FULL ACCESS

KKL Theorem for the Influence of a Set of Variables

Author: Tomasz Przybyłowski | AUTHORS INFO & AFFILIATIONS

https://doi.org/10.1137/25M174320X

Abstract.

Consider a Boolean function f on the n-dimensional hypercube, and a set of variables (indexed by) $S \subseteq \{1, 2, \dots, n\}$. The coalition influence of the variables S on a function f is the probability that after a random assignment of variables not in S, the value of f is undetermined. In this paper, we study a complementary notion, which we call the joint influence: the probability that, after a random assignment of variables not in S, the value of f is dependent on all variables in S. We show that for an arbitrary fixed d, every Boolean function f on n variables admits a d-set of joint influence at least $\frac{1}{10}\mathbf{W}^{\geq d}(f)(\frac{\log n}{n})^d$, where $\mathbf{W}^{\geq d}(f)$ is the Fourier weight of f at degrees at least d. This result is a direct generalization of the Kahn–Kalai–Linial Theorem. Further, we give an example demonstrating essential sharpness of the above bound. In our study of the joint influence we consider another notion of multibit influence recently introduced by Tal.

PDF Help

Keywords

Boolean function | influence | discrete cube

MSC codes

42C10 60E15 68R01

1. Introduction.

Consider $i \in [n] = \{1, 2, \dots, n\}$ and $f\colon \{-1,1\}^n \to \{-1,1\}$. Throughout, we consider the uniform measure μ_n on $\{-1,1\}^n$. Define the *influence* of bit i on f by

$$\mathrm{Inf}_i(f) := \mathop{\mathbb{P}}_{\mathbf{x}\sim\mu_n}(f(\mathbf{x}) \neq f(\mathbf{x}^{\oplus i})),$$

where $x^{\oplus i} = (x_1, \dots, x_{i-1}, -x_i, x_{i+1}, \dots, x_n)$.

The influence quantifies how often a given bit can sway the outcome of the function. It is a well-studied notion central to the field of Boolean analysis. Its history and various properties can be found in the textbook of O'Donnell [12]. A celebrated result, due to Kahn, Kalai, and Linial [8], says that every Boolean function admits a very influential bit.

THEOREM 1.1 (Kahn–Kalai–Linial (KKL) theorem).

There is a universal constant $C_1 > 0$ such that for every $f\colon \{-1,1\}^n \to \{-1,1\}$ there is $i \in [n]$ such that

$$\mathrm{Inf}_i(f) \geqslant C_1 \mathbb{V}\mathrm{ar}(f) \frac{\ln n}{n},$$

where $\mathbb{V}\mathrm{ar}(f) = \mathbb{V}\mathrm{ar}_{\mathbf{x}\sim\mu_n}(f(\mathbf{x}))$.

This is the best possible bound up to constant C_1 as certified by the *tribe functions* introduced by Ben-Or and Linial [2].

Several generalizations of the definition of the single-bit influence accounting for multiple bits have been considered; see [3] for discussion. The one probably most studied is the coalition influence,

PDF Help

which we will now define.

In what follows, for brevity, we associate a variable $x_i \in \{-1,1\}$ with its index i. Let $\mathbf{i} \subseteq [n]$ be a set of variables. The coalition influence $\mathrm{CInf}_{\mathbf{i}}(f)$ of $\mathbf{i}$ on f is defined as the probability that assigning values to the variables not in $\mathbf{i}$ at random, the value of f is undetermined. There has been a long line of research studying this type of influence; for the main results, see [8, 1, 5].

The focus of our paper is a counterpart notion to the coalition influence. We call it the *joint influence* $\mathrm{JInf}_{\mathbf{i}}(f)$ of $\mathbf{i}$ on f and define it as the probability that assigning values to the variables not in $\mathbf{i}$ at random, the value of f is dependent on *all* of the variables in $\mathbf{i}$.

More formally, say that $i \in \mathbf{i}$ is $\mathbf{i}$-*pivotal* for f on input x if there is an input y agreeing with x on bits $[n] \setminus \mathbf{i}$ such that $f(y) \neq f(y^{\oplus i})$. Then for nonempty $\mathbf{i}$ define

$$\mathrm{JInf}_{\mathbf{i}}(f) := \mathop{\mathbb{P}}_{\mathbf{x}\sim\mu_n}(\text{every } i \in \mathbf{i} \text{ is } \mathbf{i}\text{-pivotal for } f \text{ on input } \mathbf{x}).$$

Note that using $\mathbf{i}$-pivotality, the coalition influence satisfies

$$\mathrm{CInf}_{\mathbf{i}}(f) = \mathop{\mathbb{P}}_{\mathbf{x}\sim\mu_n}(\text{some } i \in \mathbf{i} \text{ is } \mathbf{i}\text{-pivotal for } f \text{ on input } \mathbf{x}).$$

In particular, CInf can be thought of as a *union* of influences, whereas JInf as an *intersection,* and hence the name. As a result, it follows that $\mathrm{CInf}_{\mathbf{i}}(f) \geqslant \mathrm{JInf}_{\mathbf{i}}(f)$.

Let us mention yet another generalization of the notion of single bit influence, purely analytic one, as it will turn out useful in studying the joint influence. For $\mathbf{i} \subseteq [n]$ we define the *Walsh functions* $\chi_{\mathbf{i}}\colon \{-1,1\}^n \to \mathbb{R}$ by $\chi_{\mathbf{i}}(x) := \prod_{j\in\mathbf{i}} x_j$ and $\chi_{\varnothing}(x) := 1$. The system $\{\chi_{\mathbf{i}}\colon \mathbf{i} \subseteq [n]\}$ is an orthonormal basis of $L^2(\{-1,1\}^n, \mu_n)$, which yields that every function $f\colon \{-1,1\}^n \to \mathbb{R}$ has a unique *Fourier–Walsh expansion* $f = \sum_{\mathbf{i}\subseteq[n]} \hat{f}(\mathbf{i})\chi_{\mathbf{i}}$, where $\hat{f}(\mathbf{i}) \in \mathbb{R}$. We define the *T-in* $\mathrm{Inf}_{\mathbf{i}}(f)$ by

PDF
Help

$$\mathrm{Inf}_{\mathbf{i}}(f) := \sum_{\substack{\mathbf{j}\subseteq[n]\\ \mathbf{j}\supseteq\mathbf{i}}} \hat{f}(\mathbf{j})^2.$$

It has been introduced by Tal in [14] in his study of the Fourier mass at high levels in $\mathbf{AC^0}$ Boolean circuits. However, it has been independently considered by Tanguy in [15] for two-element $\mathbf{i}$'s and in this case further studied by Oleszkiewicz in [13]. Often, when clear from the context, we will refer to it simply as the *influence.* Sometimes, we will say that $\mathrm{Inf}_{\mathbf{i}}(f)$ is a d*-degree* influence, where $d = |\mathbf{i}|$.

The T-influence, although interesting on its own, will open to us the Fourier methods to bound the joint influence via inequality $\mathrm{JInf}_{\mathbf{i}}(f) \geqslant \mathrm{Inf}_{\mathbf{i}}(f)$, proven in Proposition 3.5. Further, note that in the case of $\mathbf{i} = \{i\}$ all the three generalization of the single-bit influence are equal to $\mathrm{Inf}_i(f)$ (with the equality $\mathrm{Inf}_{\{i\}}(f) = \mathrm{Inf}_i(f)$ being folklore).

1.1. T-influence results.

Let us say a few words on what relevant to us has been proved on the T-influence.

The case of $|\mathbf{i}| = 2$ has been considered by Tanguy in [15], where he obtained a result in the spirit of the KKL Theorem. Namely, he proved that every Boolean f either admits a pair of bits with T-influence at least $C\mathbb{V}\mathrm{ar}(f)(\frac{\ln n}{n})^2$, or there is a single bit with influence at least $C(\mathbb{V}\mathrm{ar}(f)/n)^\theta$, for some universal constants $C > 0$ and $\theta \in (1/2, 1)$.

Later, Oleszkiewicz [13] extended some of Tanguy's bounds and proved that an appropriate upper bound on influences of all pairs of bits of a Boolean f implies that f is close to some affine Boolean function.

THEOREM 1.2 (Oleszkiewicz).

There exists a universal constant $C_2 > 0$ with the following property. Let $n \geqslant 2$. Assume that $f\colon \{-1,1\}^n \to \{-1,1\}$ satisfies $\mathrm{Inf}_{\{i,j\}}(f) \leqslant \alpha(\frac{\ln n}{n})^2$ for all $1 \leqslant i < j \leqslant n$ for some $\alpha > 0$. Then either

PDF Help

$$\hat{f}(\varnothing)^2 \geqslant 1 - C_2\alpha, \ \ \textit{or there exists exactly one } i \in [n] \textit{ such that } \hat{f}(\{i\})^2 \geqslant 1 - C_2\alpha$$

As Oleszkiewicz notes, the above theorem might be viewed as both a variation of the KKL Theorem and another celebrated result due to Friedgut, Kalai, and Naor [6]. For the purpose of the background of Oleszkiewicz's and our results, let us recall a generalization of the Friedgut–Kalai–Naor (FKN) Theorem due to Kindler and Safra [10] (the FKN Theorem is recovered by taking $d = 1$).

DEFINITION 1.3.

Let $d \geqslant 0$ be an integer. We say that $g\colon \{-1,1\}^n \to \{-1,1\}$ is of d-degree if, for every $\mathbf{i} \subseteq [n]$ such that $|\mathbf{i}| > d$, we have $\hat{g}(\mathbf{i}) = 0$. Equivalently, g is of d-degree if its degree as a Fourier–Walsh polynomial is at most d.

THEOREM 1.4 (Kindler–Safra).

Let d be a positive integer. If $f\colon \{-1,1\}^n \to \{-1,1\}$ satisfies $\mathbf{W}^{\geqslant d+1}(f) \leqslant \alpha$, then there exists $g\colon \{-1,1\}^n \to \{-1,1\}$ of d-degree such that

$$\|f - g\|_2^2 \leqslant C_{KS}\alpha,$$

where C_{KS} is a constant depending only on d and $\mathbf{W}^{\geqslant d+1}(f) = \sum_{\mathbf{i}\subseteq[n]:|\mathbf{i}|\geqslant d+1} \hat{f}(\mathbf{i})^2$.

In the course of the proof of Theorem 1.2, Oleszkiewicz shows that every Boolean function f admits a pair of distinct bits i, j such that $\mathrm{Inf}_{\{i,j\}}(f) \geqslant C_O\mathbf{W}^{\geqslant 2}(f)(\frac{\ln n}{n})^2$. Hence the assumption in Theorem 1.2 of small pair influences yields that $\mathbf{W}^{\geqslant 2}(f)$ is small, which, in particular, implies the assumption of Theorem 1.4 for $d = 1$. However, the stronger assumption in Theorem 1.2 yi[illegible] stronger consequence of extra $\ln n/n$ factor.

PDF
Help

1.2. Main results.

Inspired by the above, we prove the following theorem.

THEOREM 1.5.

Let $n \geqslant d \geqslant 1$ be integers. For every $f\colon \{-1,1\}^n \to \{-1,1\}$ there exists $\mathbf{i} \subseteq [n]$ of size d so that

$$\mathrm{JInf}_{\mathbf{i}}(f) \geqslant \mathrm{Inf}_{\mathbf{i}}(f) \geqslant \frac{1}{10}\mathbf{W}^{\geqslant d}(f)\left(\frac{\ln n}{n}\right)^d,$$

where $\mathbf{W}^{\geqslant d}(f) = \sum_{\mathbf{i}\subseteq[n]:|\mathbf{i}|\geqslant d} \hat{f}(\mathbf{i})^2$.

Observe that for $d = 1$ we recover the KKL Theorem, and the case $d = 2$ is the above-mentioned consequence of the proof of Theorem 1.2. We prove Theorem 1.5 in section 4 by generalizing the representation of $\mathbf{W}^{\geqslant 2}(f)$ and the hypercontractivity lemma from [13]. Recall that the first inequality holds actually for every $\mathbf{i}$ thanks to Proposition 3.5.

In section 5 we prove that Theorem 1.5 is asymptotically sharp in the following sense.

THEOREM 1.6.

For every $d \geqslant 2$ there is a constant C_H^d depending only on d, and a family of functions H_n^d such that for every $\mathbf{i} \subseteq [n]$ of size d we have

$$\mathrm{Inf}_{\mathbf{i}}(H_n^d) \leqslant \mathrm{JInf}_{\mathbf{i}}(H_n^d) \leqslant C_H^d \mathbf{W}^{\geqslant d}(H_n^d)\left(\frac{\log_2 n}{n}\right)^d.$$

The functions H_n^d, which we call the *d-hypertribes,* are a variation on the classical tribe functions due to Ben-Or and Linial [2] demonstrating the sharpness of the KKL Theorem. As was pointed out in [13], the standard tribe function cannot serve as a sharp example in our case. For example, influences of pairs of bits from the same tribe are larger than the influence of a typical pair account for this, one needs to increase the size of tribes and allow voters to belong to mult in a controlled way. These adjustments complicate the structure of the function, and the explanation why it serves as a sharp example becomes far from trivial. The construction of the d-hypertribes was suggested to us by Peter Keevash in a private communication.

Further, in sections 6 and 7 we prove the following d-degree generalization of Theorem 1.2.

PDF
Help

THEOREM 1.7.

Let $n > d \geqslant 1$. There are positive constants C_3, C_4 depending only on d such that if $f\colon \{-1,1\}^n \to \{-1,1\}$ satisfies

$$\mathrm{Inf}_{\mathbf{i}}(f) \leqslant \alpha \left(\frac{\ln n}{n}\right)^{d+1} \tag{1}$$

for all $\mathbf{i} \subseteq [n]$ of size $d+1$ for some $\alpha \in (0, C_3)$, then there exists $g\colon \{-1,1\}^n \to \{-1,1\}$ of degree d such that

$$|\hat{f}(\mathbf{j}) - \hat{g}(\mathbf{j})| \leqslant C_4 \alpha \left(\frac{\ln n}{n}\right)^{|\mathbf{j}|} \tag{2}$$

for every $\mathbf{j} \subseteq [n]$ of size at most d.

For $d = 1$ we recover Theorem 1.2. Indeed, the only 1-degree Boolean functions are the constants $-1, +1$, single-bit functions $\chi_{\{i\}} \equiv x_i$ and their negations. Depending which of these g is, from (2) we can recover the relevant alternative of Theorem 1.2.

We will see in section 3 that $\mathrm{Inf}_{\mathbf{i}}(f) = \|\partial_{\mathbf{i}} f\|_2^2$, where $\partial_{\mathbf{i}} f$ is a (discrete) d-order partial derivative of f. From this point of view, Theorem 1.7 is a Boolean variant of a theorem from real analysis stating roughly that if all the $(d+1)$-order partial derivatives of $f\colon [0,1]^n \to \mathbb{R}$ are small, then f is close to some polynomial of degree at most d.

Similarly to Theorem 1.2, Theorem 1.7 might be viewed as a variation on the Kindler–Safra Theorem 1.4. And exactly like there, the assumption (1) implies the Kindler and Safra assumption: we have $\mathbf{W}^{\geqslant d+1}(f) \leqslant 10\alpha$. However, Theorem 1.7 tells us that this stronger assumption yields a l
approximation.

Further, if we additionally assume that f has small coefficients of low frequencies, then we obtain a strictly better bound on the distance $\|f - g\|_2^2$ compared to the Kindler–Safra Theorem. We prove this at the end of section 7.

PDF
Help

COROLLARY 1.8.

Let $n > d \geqslant 1$. Assume that $f\colon \{-1,1\}^n \to \{-1,1\}$ satisfies (1) *and that for some $l \in [d]$ we have $|\hat{f}(\mathbf{i})| \leqslant 2^{-d}$ for all $\mathbf{i} \subseteq [n]$ of size at most $l-1$. Then there exists a function $g\colon \{-1,1\}^n \to \{-1,1\}$ of d-degree such that*

$$\|f-g\|_2^2 \leqslant C_5 \alpha \left(\frac{\ln n}{n}\right)^l,$$

where C_5 is a constant depending only on d.

2. Notations.

Throughout this paper, n is an arbitrary positive integer (possibly further constrained). Let $d \in \mathbb{Z}_{\geqslant 0}$. We use the standard notation $[n] = \{1, 2, \dots, n\}$ and the less standard notation

$$[n]_d := \{A \subseteq [n]\colon |A| = d\}, \quad [n]_{<d} := \{A \subseteq [n]\colon |A| < d\}.$$

We define $[n]_{\leqslant d}$, $[n]_{>d}$, and $[n]_{\geqslant d}$ in an analogous manner.

We write $\mathrm{Im}(f)$ for the image of a function f, and refer to functions with $\mathrm{Im}(f) \subseteq \{-1,1\}$ as *Boolean* functions. Plain letters i, j, k usually denote indices of bits of the input, i.e., elements of $[n]$, whereas bold letters $\mathbf{i}, \mathbf{j}, \mathbf{k}$ refer to index sets, i.e., subsets of $[n]$. For brevity, we will associate a variable $x_i \in \{-1,1\}$ with its index i, and so we will say, e.g., *variables* $\mathbf{i}$ instead of *variables* $\{x_i\colon i \in \mathbf{i}\}$. Sometimes, for clarity, we omit brackets and write, e.g., $\mathbf{i} \setminus i$ instead of $\mathbf{i} \setminus \{i\}$. The name *$d$-set* will stand for a d element subset of $[n]$.

For an input $x = (x_1, \dots, x_n), i \in [n]$, and $\varepsilon \in \{-1,1\}$ we write

$$x^{\oplus i} = (x_1, \dots, x_{i-1}, -x_i, x_{i+1}, \dots, x_n),$$
$$x^{i \to \varepsilon} = (x_1, \dots, x_{i-1}, \varepsilon, x_{i+1}, \dots, x_n),$$

PDF
Help

and further, for $\mathbf{i} \subseteq [n]$ and $y \in \{-1,1\}^{\mathbf{i}}$ define $x^{\mathbf{i}\to y} \in \{-1,1\}^n$ by

$$x_j^{\mathbf{i}\to y} = \begin{cases} x_j & \text{if } j \in [n] \setminus \mathbf{i}, \\ y_j & \text{if } j \in \mathbf{i}. \end{cases}$$

3. Basics of harmonic analysis.

In this section we introduce the standard language of the Boolean analysis. The basic Fourier–Walsh notions are recalled and we repeat the definition of influences in a broader context. Finally, we state some results regarding *hypercontractivity,* which we will use extensively in the proofs of both main theorems. Many of these notions (with exception of the multibit influences) are explained in greater detail in O'Donnell's book [12].

Let $\mu_n = (\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes n}$ be the uniform measure on $\{-1,1\}^n$. Let us consider the linear space $L^2(\{-1,1\}^n, \mu_n)$ of all functions $f\colon \{-1,1\}^n \to \mathbb{R}$ equipped with the standard inner product and norms given by

$$\langle f, g\rangle := \frac{1}{2^n} \sum_{x\in\{-1,1\}^n} f(x)g(x) = \mathbb{E}[f(\mathbf{x})g(\mathbf{x})] = \mathbb{E}[fg] \quad \text{and}$$
$$\|f\|_2 := \sqrt{\langle f, f\rangle} = \sqrt{\mathbb{E}[f(\mathbf{x})^2]} = \sqrt{\mathbb{E}f^2}.$$

The probabilities $\mathbb{P}$ and expectations $\mathbb{E}$ are always taken with respect to the measure μ_n. The bold letter $\mathbf{x}$ will then denote a uniform random variable with distribution μ_n. Note that, due to the product structure of μ_n, the set of single bits $\mathbf{x}_i$ of $\mathbf{x}$ is a family of independent and identically distributed random variables, each having distribution $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$. Often, when it is clear from the context, we omit the letter $\mathbf{x}$ in probabilities or expectations.

For $\mathbf{i} \subseteq [n]$ we define the *Walsh functions* $\chi_{\mathbf{i}}\colon \{-1,1\}^n \to \mathbb{R}$ by $\chi_{\mathbf{i}}(x) := \prod_{j\in\mathbf{i}} x_j$ and $:= 1$. The system $\{\chi_{\mathbf{i}}\colon \mathbf{i} \subseteq [n]\}$ is an orthonormal basis of $L^2(\{-1,1\}^n, \mu_n)$, which yie every function $f\colon \{-1,1\}^n \to \mathbb{R}$ has a unique *Fourier–Walsh expansion*

PDF Help

$$f = \sum_{\mathbf{i} \subseteq [n]} \hat{f}(\mathbf{i}) \chi_{\mathbf{i}},$$

where we call $\hat{f}(\mathbf{i}) := \langle f, \chi_{\mathbf{i}} \rangle \in \mathbb{R}$ a *Fourier–Walsh coefficient.*

In this setting, every $f, g\colon \{-1,1\}^n \to \mathbb{R}$ satisfies the *Plancherel* and *Parseval identities*

$$\langle f, g \rangle = \sum_{\mathbf{k} \subseteq [n]} \hat{f}(\mathbf{k}) \hat{g}(\mathbf{k}) \quad \text{and} \quad \|f\|_2^2 = \sum_{\mathbf{k} \subseteq [n]} \hat{f}(\mathbf{k})^2.$$

Further, note that we have $\mathbb{E} f = \langle f, 1 \rangle = \langle f, \chi_\varnothing \rangle = \hat{f}(\varnothing)$. Moreover, if f is Boolean (so $\operatorname{Im} f \subseteq \{-1,1\}$) we have $\|f\|_2^2 = \sum_{\mathbf{i} \subseteq [n]} \hat{f}(\mathbf{i})^2 = 1$ and so in this case $|\hat{f}(\mathbf{i})| \leqslant 1$ for every $\mathbf{i} \subseteq [n]$.

Also, for $f\colon \{-1,1\}^n \to \mathbb{R}$ we define

$$\mathbf{W}^{\geqslant d}(f) := \sum_{\mathbf{i} \in [n]_{\geqslant d}} \hat{f}(\mathbf{i})^2 \quad \text{and} \quad \mathbf{W}^{=d}(f) := \sum_{\mathbf{i} \in [n]_d} \hat{f}(\mathbf{i})^2.$$

We will say that f is of *d-degree* if for every $\mathbf{i} \in [n]_{>d}$ we have $\hat{f}(\mathbf{i}) = 0$. Equivalently, f is of d-degree if its degree as a Fourier–Walsh polynomial is at most d.

3.1. Derivative.

In the following section we introduce the notion of the discrete partial derivatives. We remark that observations regarding partial derivatives (Definition 3.1 and Proposition 3.2 below) were already made in [14].

Let $f\colon \{-1,1\}^n \to \mathbb{R}, i \in [n]$. Define the *$i$th partial derivative* $\partial_i f\colon \{-1,1\}^n \to \mathbb{R}$ by[1]

$$\partial_i f(x) := \frac{f(x^{i \to +1}) - f(x^{i \to -1})}{2}.$$

PDF Help

By linearity, one can easily get a useful Fourier–Walsh representation

$$\partial_i f = \sum_{\substack{\mathbf{j}\subseteq[n]\\ \mathbf{j}\ni i}} \hat{f}(\mathbf{j})\chi_{\mathbf{j}\setminus\{i\}}. \tag{3}$$

The derivative operators may be composed. The order of composition does not matter due to the Fourier–Walsh representation (3). In consequence, we can give the following general definition.

DEFINITION 3.1.

Let $f\colon\{-1,1\}^n\to\mathbb{R}$ *and* $\mathbf{i}=\{i_1,i_2,\dots,i_k\}\subseteq[n]$. *We define the* $\mathbf{i}$th partial derivative $\partial_{\mathbf{i}}f\colon\{-1,1\}^n\to\mathbb{R}$ *by*

$$\partial_{\mathbf{i}}f := \partial_{i_1}\circ\partial_{i_2}\circ\cdots\circ\partial_{i_k}f = \sum_{\substack{\mathbf{j}\subseteq[n]\\ \mathbf{j}\supseteq\mathbf{i}}} \hat{f}(\mathbf{j})\chi_{\mathbf{j}\setminus\mathbf{i}} \qquad \text{and} \qquad \partial_{\varnothing}f := f.$$

A trivial induction proves that the definition of the single-bit derivative extends to a multibit counterpart in the following way.

PROPOSITION 3.2.

Let $f\colon\{-1,1\}^n\to\mathbb{R}$ *and* $\mathbf{i}\subseteq[n]$ *be nonempty. Then*

$$\partial_{\mathbf{i}}f(x) = \frac{1}{2^{|\mathbf{i}|}}\sum_{y\in\{-1,1\}^{\mathbf{i}}}\left(f(x^{\mathbf{i}\to y})\cdot\prod_{j\in\mathbf{i}}y_j\right). \tag{4}$$

COROLLARY 3.3.

Let $f\colon\{-1,1\}^n\to\{-1,1\}$ *and* $\mathbf{i}\subseteq[n]$ *be nonempty. Then*

PDF Help

$$\operatorname{Im}(\partial_{\mathbf{i}} f) \subseteq \mathbb{Z}/2^{|\mathbf{i}|-1} \cap [-1,1].$$

Proof.

The definition of the single-bit derivative, or Proposition 3.2, easily yields $\operatorname{Im}(\partial_{\mathbf{i}} f) \subseteq [-1,1]$. For the other containment, observe that after application of Proposition 3.2, all we need to prove is that the sum in (4) is an even integer. This is indeed the case as it is a sum of $2^{|\mathbf{i}|}$ odd integers. □

3.2. Influence.

In the following section we will connect the influence definitions from the introduction with the notion of derivative, and prove some useful bounds between them.

First, let us connect the definition of influence of a single-bit with the Fourier–Walsh coefficients. Observe that $f(x) \neq f(x^{\oplus i})$ occurs if and only if $\partial_i f(x) \neq 0$, which occurs if and only if $|\partial_i f(x)| = 1$. Therefore, by Parseval's identity

$$\operatorname{Inf}_i(f) = \mathbb{P}(\partial_i f(\mathbf{x}) \neq 0) = \mathbb{E}[(\partial_i f(\mathbf{x}))^2] = \|\partial_i f\|_2^2 = \sum_{\substack{\mathbf{j} \subseteq [n] \\ \mathbf{j} \ni i}} \hat{f}(\mathbf{j})^2.$$

In particular, as we promised in the introduction, $\operatorname{Inf}_i(f)$ is equal to the T-influence $\operatorname{Inf}_{\{i\}}(f)$.

The definition of the multibit derivative yields an analogous representation of the T-influence. Let $f\colon \{-1,1\}^n \to \mathbb{R}$ and $\mathbf{i} \subseteq [n]$. Again, by Parseval's identity we have

$$\operatorname{Inf}_{\mathbf{i}}(f) = \sum_{\substack{\mathbf{j} \subseteq [n] \\ \mathbf{j} \supseteq \mathbf{i}}} \hat{f}(\mathbf{j})^2 = \|\partial_{\mathbf{i}} f\|_2^2.$$

PDF Help

PROPOSITION 3.4.

Let $f\colon \{-1,1\}^n \to \{-1,1\}$, $\mathbf{i} \subseteq [n]$, and $r = |\mathbf{i}| > 0$. Then

i. $\mathbb{P}(\partial_{\mathbf{i}} f \neq 0) \geqslant \mathrm{Inf}_{\mathbf{i}}(f) \geqslant \frac{1}{2^{2r-2}} \mathbb{P}(\partial_{\mathbf{i}} f \neq 0)$;
ii. $\mathrm{Inf}_{\mathbf{i}}(f) \geqslant \frac{1}{2^{r-1}} |\hat{f}(\mathbf{i})|$.

Proof.

By Corollary 3.3 we have $(\partial_{\mathbf{i}} f)^2 \in [\mathbb{1}_{\{\partial_{\mathbf{i}} f \neq 0\}} / 2^{2r-2}, \mathbb{1}_{\{\partial_{\mathbf{i}} f \neq 0\}}]$, so taking expectation yields (i).

For (ii), again, by Corollary 3.3 we obtain

$$\mathrm{Inf}_{\mathbf{i}}(f) = \mathbb{E}[|\partial_{\mathbf{i}} f|^2] \geqslant \mathbb{E}\left[\frac{1}{2^{r-1}} |\partial_{\mathbf{i}} f|\right] \geqslant \frac{1}{2^{r-1}} |\mathbb{E}[\partial_{\mathbf{i}} f]| = \frac{1}{2^{r-1}} |\hat{f}(\mathbf{i})|.$$

□

PROPOSITION 3.5.

Let $f\colon \{-1,1\}^n \to \{-1,1\}$, and let a nonempty $\mathbf{i} \subseteq [n]$ satisfy $\partial_{\mathbf{i}} f(x) \neq 0$ for some $x \in \{-1,1\}^n$. Then every $i \in \mathbf{i}$ is $\mathbf{i}$-pivotal for f on x. In particular, $\mathrm{JInf}_{\mathbf{i}}(f) \geqslant \mathrm{Inf}_{\mathbf{i}}(f)$.

Proof.

Take an arbitrary $i \in \mathbf{i}$, and denote by $y^+ = y^{i \to +1}$ and $y^- = y^{i \to -1}$. By Proposition 3.2 we have

$$\begin{aligned} \partial_{\mathbf{i}} f(x) = (\partial_i(\partial_{\mathbf{i}\setminus i} f))(x) &= \frac{\partial_{\mathbf{i}\setminus i} f(x^{i \to +1}) - \partial_{\mathbf{i}\setminus i} f(x^{i \to -1})}{2} \\ &= \frac{1}{2^{|\mathbf{i}|}} \sum_{y \in \{-1,1\}^{\mathbf{i}\setminus i}} \left(f(x^{\mathbf{i} \to y^+}) - f(x^{\mathbf{i} \to y^-})\right) \prod_{j \in \mathbf{i}\setminus i} y_j. \end{aligned}$$

Since $\partial_{\mathbf{i}} f(x) \neq 0$, there exists a nonzero summand in the above sum. As a result, i is $\mathbf{i}$-pivotal on x, which holds for every $i \in \mathbf{i}$.

For the inequality part, the above implies $\mathrm{JInf}_{\mathbf{i}}(f) \geqslant \mathbb{P}(\partial_{\mathbf{i}} f \neq 0)$ and so by Proposition 3.4 we get $\mathrm{JInf}_{\mathbf{i}}(f) \geqslant \mathrm{Inf}_{\mathbf{i}}(f)$. □

PDF Help

Remark 3.6.

For $\mathbf{i}$ of size 1 or 2 one can check that the first part of the proposition is actually an equivalence, in which case $\mathbf{JInf}_{\mathbf{i}}(f) = \mathbb{P}(\partial_{\mathbf{i}} f \neq 0)$.

Let us introduce for brevity two more pieces of notation.

Definition 3.7.

Let $f\colon \{-1,1\}^n \to \mathbb{R}$, and let $d \leqslant n$ be a positive integer. We define

$$\mathbf{Inf}(f) := \sum_{i\in[n]} \mathbf{Inf}_i(f) \quad \textit{and} \quad \mathbf{MaxInf}_d(f) := \max_{\mathbf{i}\in[n]_d} \mathbf{Inf}_{\mathbf{i}}(f).$$

3.3. Hypercontraction and the log-Sobolev inequality.

The key element of the proofs of the KKL, FKN and Kindler–Safra Theorems is the celebrated hypercontractivity theorem by Bonami [4], presented below. It is also the central ingredient of our proofs of Theorems 1.5 and 1.7.

Definition 3.8.

Let $f\colon \{-1,1\}^n \to \mathbb{R}$ and $t \geqslant 0$. We define $P_t f\colon \{-1,1\}^n \to \mathbb{R}$ by

$$P_t f := \sum_{\mathbf{i}\subseteq[n]} e^{-|\mathbf{i}|t} \hat{f}(\mathbf{i}) \chi_{\mathbf{i}}.$$

The family of operators $(P_t)_{t\geqslant 0}$ is often called the heat semigroup *on the Boolean cube.*

Theorem 3.9 (see [4]).

Let $f\colon \{-1,1\}^n \to \mathbb{R}$ and $t \geqslant 0$. Then

PDF
Help

$$\|P_t f\|_2 \leqslant \|f\|_{1+e^{-2t}}.$$

The second important fact we make use of is Lemma 3.10—a weak functional version of the edge isoperimetric inequality—which is a consequence of the classical log-Sobolev inequality. We will use it in proving Theorem 1.7.

LEMMA 3.10.

If $f\colon \{-1,1\}^n \to \{0,1\}$, *then* $\mathrm{Inf}(f) \geqslant \frac{1}{2}\mathbb{E}f \cdot \ln(1/\mathbb{E}f)$.

For the proof and background of the lemma we refer the reader to section 5 of [13].

4. Proof of Theorem 1.5.

In this section we prove Theorem 1.5. Our proof proceeds via two propositions. The first, Proposition 4.1, provides an integral representation for the higher order variance $\mathbf{W}^{\geqslant d}(f)$ via the heat semigroup operator P_t. The second, Proposition 4.2, uses hypercontractivity and gives a bound on integrals of the representation of the higher order variance in terms of influences. These propositions are generalizations of Lemmas 3.1 and 3.2 from [13]. The proofs presented proceed in a fairly analogous manner to those in [13].

PROPOSITION 4.1.

Let $n \geqslant d \geqslant 1$ *be integers. For every* $f\colon \{-1,1\}^n \to \mathbb{R}$ *we have*

$$\mathbf{W}^{\geqslant d}(f) = \sum_{\mathbf{i}\in[n]_{\geqslant d}} \hat{f}(\mathbf{i})^2 = 2d \sum_{\mathbf{i}\in[n]_d} \int_0^\infty (e^{2t}-1)^{d-1} e^{-2dt} \|P_t \partial_{\mathbf{i}} f\|_2^2 dt.$$

Proof.

Since for $\mathbf{i} \in [n]_d$ we have

PDF Help

$$P_t\partial_{\mathbf{i}} f = P_t\left(\sum_{\substack{\mathbf{k}\subseteq[n]\\ \mathbf{i}\subseteq\mathbf{k}}} \hat{f}(\mathbf{k})\chi_{\mathbf{k}\setminus\mathbf{i}}\right) = \sum_{\substack{\mathbf{k}\subseteq[n]\\ \mathbf{i}\subseteq\mathbf{k}}} \mathrm{e}^{-(|\mathbf{k}|-d)t}\hat{f}(\mathbf{k})\chi_{\mathbf{k}\setminus\mathbf{i}},$$

by the Parseval identity we obtain

$$\begin{aligned}\sum_{\mathbf{i}\in[n]_d} \mathrm{e}^{-2dt}\|P_t\partial_{\mathbf{i}} f\|_2^2 &= \sum_{\mathbf{i}\in[n]_d} \mathrm{e}^{-2dt}\sum_{\substack{\mathbf{k}\subseteq[n]\\ \mathbf{i}\subseteq\mathbf{k}}} \mathrm{e}^{-2(|\mathbf{k}|-d)t}\hat{f}(\mathbf{k})^2 = \sum_{\mathbf{i}\in[n]_d}\sum_{\substack{\mathbf{k}\subseteq[n]\\ \mathbf{i}\subseteq\mathbf{k}}} \mathrm{e}^{-2|\mathbf{k}|t}\hat{f}(\mathbf{k})^2\\ &= \sum_{\mathbf{k}\in[n]_{\geqslant d}}\sum_{\substack{\mathbf{i}\in[n]_d\\ \mathbf{i}\subseteq\mathbf{k}}} \mathrm{e}^{-2|\mathbf{k}|t}\hat{f}(\mathbf{k})^2 = \sum_{\mathbf{k}\in[n]_{\geqslant d}}\binom{|\mathbf{k}|}{d}\mathrm{e}^{-2|\mathbf{k}|t}\hat{f}(\mathbf{k})^2.\end{aligned}$$

Therefore,

$$\sum_{\mathbf{i}\in[n]_d}\int_0^\infty (\mathrm{e}^{2t}-1)^{d-1}\mathrm{e}^{-2dt}\|P_t\partial_{\mathbf{i}} f\|_2^2\mathrm{d}t = \sum_{\mathbf{k}\in[n]_{\geqslant d}}\binom{|\mathbf{k}|}{d}\hat{f}(\mathbf{k})^2\int_0^\infty (\mathrm{e}^{2t}-1)^{d-1}\mathrm{e}^{-2|\mathbf{k}|t}\mathrm{d}t.$$

Substituting $u=\mathrm{e}^{-2t}, u\in(0,1]$, i.e., $t=-\frac{1}{2}\ln u$ with $\left|\frac{\mathrm{d}t}{\mathrm{d}u}\right| = \frac{1}{2u}$, we get

$$\begin{aligned}\int_0^\infty (\mathrm{e}^{2t}-1)^{d-1}\mathrm{e}^{-2|\mathbf{k}|t}\mathrm{d}t &= \int_0^1\left(\frac{1}{u}-1\right)^{d-1}u^{|\mathbf{k}|}\cdot\frac{\mathrm{d}u}{2u} = \frac{1}{2}\int_0^1(1-u)^{d-1}u^{|\mathbf{k}|-d}\mathrm{d}u\\ &= \frac{1}{2}\mathrm{B}(d,|\mathbf{k}|-d+1) = \frac{1}{2d\cdot\binom{|\mathbf{k}|}{d}},\end{aligned}$$

which yields the statement. □

PROPOSITION 4.2.

Let n, d, l be integers such that $d \geqslant l \geqslant 1$. Assume that $f\colon \{-1,1\}^n \to \mathbb{R}$ and some $\mathbf{i} \subseteq [n]$ satisfy $\mathrm{Im}(\partial_{\mathbf{i}} f) \subseteq \frac{1}{2^{d-1}}\mathbb{Z}$ and $\mathrm{Inf}_{\mathbf{i}}(f) \leqslant 1$. Then

$$\int_0^\infty (e^{2t}-1)^{l-1} e^{-2lt} \|P_t \partial_{\mathbf{i}} f\|^2 dt \leqslant (l-1)! 4^{d-1} \cdot \frac{\mathrm{Inf}_{\mathbf{i}}(f)}{\ln^l(1/\mathrm{Inf}_{\mathbf{i}}(f))}.$$

Remark 4.3.

When $\mathrm{Inf}_{\mathbf{i}}(f) \in \{0,1\}$ we let $1 \cdot (\ln 1/1)^{-l} = +\infty$ and $0 \cdot (\ln 1/0)^{-l} = 0$.

Remark 4.4.

The seemingly artificial assumption that $\mathrm{Im}(\partial_{\mathbf{i}} f) \subseteq \frac{1}{2^{d-1}}\mathbb{Z}$ is satisfied already for f Boolean and $\mathbf{i} \in [n]_d$, due to Corollary 3.3. The assumption takes such a form since it will also be used in the proof of Theorem 1.7, where a slightly more general form is needed.

Proof.

If $\mathrm{Inf}_{\mathbf{i}}(f) = 1$, then the right-hand side equals $+\infty$. If $\mathrm{Inf}_{\mathbf{i}}(f) = 0$, then $\partial_{\mathbf{i}} f \equiv 0$, so the integral is equal to 0. Therefore, in both cases the statement is true. Henceforth, we will assume that $\mathrm{Inf}_{\mathbf{i}}(f) \in (0,1)$.

The Bonami hypercontractivity bound applied to the function $\partial_{\mathbf{i}} f$ gives that

$$\|P_t \partial_{\mathbf{i}} f\|_2^2 \leqslant \|\partial_{\mathbf{i}} f\|_{1+\mathrm{e}^{-2t}}^2 = \left(\mathbb{E}[|\partial_{\mathbf{i}} f|^{1+\mathrm{e}^{-2t}}]\right)^{\frac{2}{1+\mathrm{e}^{-2t}}}. \tag{5}$$

Observe that for every $v \in \mathrm{Im}(\partial_{\mathbf{i}} f) \subseteq \frac{1}{2^{d-1}}\mathbb{Z}$, we have

$$|v|^{1+\mathrm{e}^{-2t}} \leqslant 2^{(d-1)(1-\mathrm{e}^{-2t})} \cdot v^2.$$

This yields that

PDF Help

(6) $$\mathbb{E}[|\partial_{\mathbf{i}} f|^{1+\mathrm{e}^{-2t}}] \leqslant 2^{(d-1)(1-\mathrm{e}^{-2t})}\mathbb{E}[(\partial_{\mathbf{i}} f)^2] = 2^{(d-1)(1-\mathrm{e}^{-2t})} \cdot \mathrm{Inf}_{\mathbf{i}}(f).$$

Due to (5) and (6), we obtain

(7) $$\|P_t \partial_{\mathbf{i}} f\|_2^2 \leqslant (2^{2(d-1)})^{\frac{1-\mathrm{e}^{-2t}}{1+\mathrm{e}^{-2t}}} \cdot \mathrm{Inf}_{\mathbf{i}}(f)^{1+\frac{1-\mathrm{e}^{-2t}}{1+\mathrm{e}^{-2t}}} = 2^{2(d-1)u} \cdot \mathrm{Inf}_{\mathbf{i}}(f)^{1+u},$$

where we substituted $u = \frac{1-\mathrm{e}^{-2t}}{1+\mathrm{e}^{-2t}}$, i.e., $t = \frac{\ln(1+u)-\ln(1-u)}{2}$. As $t \in [0,\infty)$, we have $u \in [0,1)$ and $\frac{\mathrm{d}t}{\mathrm{d}u} = \frac{1}{(1-u)(1+u)}$. Using the bound (7) we get

$$\begin{aligned}\int_0^\infty (\mathrm{e}^{2t}-1)^{l-1}\mathrm{e}^{-2lt}\|P_t\partial_{\mathbf{i}} f\|^2 \mathrm{d}t &\leqslant \int_0^1 \left(\frac{2u}{1-u}\right)^{l-1}\left(\frac{1-u}{1+u}\right)^{l} 2^{2(d-1)u}\frac{\mathrm{Inf}_{\mathbf{i}}(f)^{1+u}}{(1+u)(1-u}\\ &\leqslant 4^{d-1}\,\mathrm{Inf}_{\mathbf{i}}(f)\int_0^1 \left(\frac{2^u}{1+u}\right)^{l+1} u^{l-1}\,\mathrm{Inf}_{\mathbf{i}}(f)^u \mathrm{d}u\\ &\leqslant 4^{d-1}\,\mathrm{Inf}_{\mathbf{i}}(f)\int_0^\infty u^{l-1}\,\mathrm{Inf}_{\mathbf{i}}(f)^u \mathrm{d}u,\end{aligned}$$

where we used the bounds $2^{l-1+2(d-1)u} \leqslant 2^{2d-2+(l+1)u}$ and $2^u \leqslant 1+u$ for $u \in [0,1]$.

Finally,

$$\begin{aligned}\int_0^\infty u^{l-1}\,\mathrm{Inf}_{\mathbf{i}}(f)^u \mathrm{d}u &= \int_0^\infty u^{l-1}\mathrm{e}^{u\ln(\mathrm{Inf}_{\mathbf{i}}(f))}\mathrm{d}u = \frac{1}{\ln^l(1/\mathrm{Inf}_{\mathbf{i}}(f))}\int_0^\infty t^{l-1}\mathrm{e}^{-t}\mathrm{d}t\\ &= \frac{(l-1)!}{\ln^l(1/\mathrm{Inf}_{\mathbf{i}}(f))},\end{aligned}$$

which finishes the proof. □

We are ready to prove Theorem 1.5.

PDF Help

Proof of Theorem 1.5.

For $n = 1$ the statement is trivial. Therefore, let us assume that $n \geqslant 2$.

Observe that for $\mathbf{i} \in [n]_d$ we have $\mathrm{Im}(\partial_{\mathbf{i}} f) \subseteq \mathbb{Z}/2^{d-1}$. Hence, by Propositions 4.1 and 4.2 for $l = d$ and all $\mathbf{i} \in [n]_d$ we obtain

$$\mathbf{W}^{\geqslant d}(f) = 2d \sum_{\mathbf{i}\in[n]_d} \int_0^\infty (\mathrm{e}^{2t} - 1)^{d-1} \mathrm{e}^{-2dt} \|P_t \partial_{\mathbf{i}} f\|_2^2 \mathrm{d}t \leqslant d! 2^{2d-1} \sum_{\mathbf{i}\in[n]_d} \frac{\mathrm{Inf}_{\mathbf{i}}(f)}{\ln^d(1/\mathrm{Inf}_{\mathbf{i}}(f))}.$$

Since the function $t \mapsto t(\ln(1/t))^{-d}$ is increasing on [0,1], we get

$$\mathbf{W}^{\geqslant d}(f) \leqslant d! 2^{2d-1} \binom{n}{d} \frac{\mathrm{MaxInf}_d(f)}{\ln^d(1/\mathrm{MaxInf}_d(f))} \leqslant 2^{2d-1} n^d \frac{\mathrm{MaxInf}_d(f)}{\ln^d(1/\mathrm{MaxInf}_d(f))}. \tag{8}$$

Assume that $\mathrm{MaxInf}_d(f) \leqslant \frac{1}{n^{d/2}}$. The function $t \mapsto (\ln(1/t))^{-d}$ is increasing on [0,1], so by (8) we obtain

$$\begin{aligned}\mathbf{W}^{\geqslant d}(f) &\leqslant 2^{2d-1} n^d \cdot \frac{\mathrm{MaxInf}_d(f)}{((d/2)\ln n)^d} = \frac{2^{3d-1}}{d^d} \cdot \Big(\frac{n}{\ln n}\Big)^d \mathrm{MaxInf}_d(f) \\ &\leqslant 10 \Big(\frac{n}{\ln n}\Big)^d \mathrm{MaxInf}_d(f),\end{aligned}$$

and so the proof in this case is done.

On the other hand, if $\mathrm{MaxInf}_d(f) > \frac{1}{n^{d/2}}$, then since $\ln n \leqslant \sqrt{n}$ for all $n \in \mathbb{Z}_+$, we ha

$$\mathrm{MaxInf}_d(f) > \frac{1}{n^{d/2}} \geqslant \Big(\frac{\ln n}{n}\Big)^d \geqslant \mathbf{W}^{\geqslant d}(f) \cdot \Big(\frac{\ln n}{n}\Big)^d,$$

PDF
Help

as $\mathbf{W}^{\geqslant d}(f) \leqslant \|f\|_2^2 = 1$, which finishes the proof. □

5. The hypertribe example.

In the following section, we show the essential sharpness of the lower-bound in Theorem 1.5. More precisely, for every $d \geqslant 2$ we construct a family of *hypertribe* functions H_n^d satisfying $\mathrm{MaxInf}_d(H_n^d) \leqslant C\mathbf{W}^{\geqslant d}(H_n^d)(\frac{\log n}{n})^d$ for every $n \geqslant d$ and some constant $C = C(d) > 0$. The construction we present was suggested to us by Peter Keevash.

Let us elaborate on some quantities behind the construction and why its 1-degree prototype does not work.

In the case of the KKL Theorem, the sharpness is provided by the (standard) tribe functions due to Ben-Or and Linial [2]. For the construction, one splits n voters (bits) into disjoint tribes of sizes roughly $\log(\frac{n}{\log n})$ and declares the outcome (i.e., output of the function) to be $+1$ if there is a tribe such that all its members voted $+1$, and to be -1, if there is no such tribe.

The standard tribe functions, as noted in [13], do not give sharpness in Theorem 1.5 already for $d = 2$. In this case, although the influences of pairs of voters from distinct tribes are of the (desired) order $(\frac{\log n}{n})^2$, the influences of pairs of voters from the same tribe are of order $\frac{\log n}{n}$. If one accounts for that, and changes the sizes of the tribes to $2\log(\frac{n}{\log n})$, the variance (and thus also $\mathbf{W}^{\geqslant 2}$) drops down from constant to $O(\frac{\log n}{n})$ magnitude, which is again not enough to get sharpness.

In order to make up for this drop, we need to increase the number of tribes. This comes at the cost of tribes no longer being disjoint, which in turn may increase the order of magnitude of the influences. To keep them of the desired small order, we need to guarantee that for every pair of voters (or, more generally, every d-tuple) they are not simultaneously members of too many tribes. Fortunately, the Kuzjurin result (the second part of Theorem 1 in [11]) regarding existence of certain asymptotically good packings guarantees the existence of the desired functions.

We start by stating Theorem 5.1, a special case of Theorem 1 from [11]. Based on this, in Example 5.2 we construct the hypertribe functions H_n^d. In Proposition 5.3, we state some basic quantitative observations regarding the structure of H_n^d. In Propositions 5.5, 5.8, and 5.9 we prove boun
influences of H_n^d and the probabilities of the outcomes ± 1. Finally, in Theorem 5.10 we co
proven bounds and show that the hypertribes indeed asymptotically attain the lower bound from Theorem 1.5.

PDF
Help

THEOREM 5.1 (see [11]).

Fix $d \geqslant 2$. Assume that $k(n)$ satisfies $k(n) \overset{n\to\infty}{\longrightarrow} \infty$ and $k(n) \leqslant c\sqrt{n}$ for some positive constant $c < 1$ and sufficiently large n. Then there exists $N > 0$ such that for every $n > N$ there is a family $\mathcal{F}_n$ of $k(n)$-element subsets of $[n]$ such that

(∗) *for at least half of the d-sets $\mathbf{i} \in [n]_d$ there is a set $A_{\mathbf{i}} \in \mathcal{F}_n$ such that $\mathbf{i} \subseteq A_{\mathbf{i}}$;*

(∗∗) *for every d-set $\mathbf{i} \in [n]_d$ there is at most one set $A \in \mathcal{F}_n$ such that $\mathbf{i} \subseteq A$.*

The family $\mathcal{F}_n$ can be understood as a *packing* of d-set-disjoint $k(n)$-sets. The above theorem asserts that this can be done in such a way that at least half of the d-sets are covered.

Example 5.2 (d-hypertribes).

Fix $d \geqslant 2$. Let $k(n) = d \log_2 \frac{n}{\log_2 n}$, which is an integer when $n = 2^{2^a}$ and $a \in \mathbb{Z}_+$. For sufficiently large n we clearly have $k(n) \leqslant \frac{1}{2}\sqrt{n}$ and $k(n) \overset{n\to\infty}{\longrightarrow} \infty$. Choose $n > N$ so that $k(n)$ is integer. Then, by Theorem 5.1, we get a family $\mathcal{F}_n$ of $k(n)$-element subsets of $[n]$ satisfying both (∗) and (∗∗).

Let us treat $+1$ and -1 as the Boolean *true* and *false,* respectively. We construct the d-hypertribe function $H_n^d \colon \{-1,1\}^n \to \{-1,1\}$ as follows:

$$H_n^d(x) := \bigvee_{A \in \mathcal{F}_n} \bigwedge_{i \in A} x_i = \mathrm{OR}\left(\mathrm{AND}(x_1^1, \ldots, x_{k(n)}^1), \ldots, \mathrm{AND}(x_1^t, \ldots, x_{k(n)}^t)\right),$$

where $t := |\mathcal{F}_n|$ and $x_i^j \in A_j \in \mathcal{F}_n$. In other words, $H_n^d(x)$ is the outcome of (some special kind of) voting of t tribes, each consisting of $k(n)$ members. If there is any tribe that unanimously votes for *true,* then the outcome of H_n^d is *true.* Otherwise, the outcome of H_n^d is *false.* Tribes are not necessarily disjoint and each voter has to send the same vote in all the tribes to which they belong.

From now on we simply write k and H_n instead of $k(n)$ and H_n^d.

Proposition 5.3.

The family $\mathcal{F}_n$ constructed in Example 5.2 satisfies the following properties.

(P1) *For every $\mathbf{i} \in [n]_{\leqslant d}$,*

PDF
Help

$$\left|\{A \in \mathcal{F}_n \colon \mathbf{i} \subseteq A\}\right| \leqslant \left(\frac{n}{\log_2 n}\right)^{d-|\mathbf{i}|}.$$

(P2) *There exists a constant $c_1 > 0$ depending only on d such that*

$$c_1 \leqslant t \cdot 2^{-k} \leqslant 1.$$

(P3) *Every two distinct tribes $A, B \in \mathcal{F}_n$ have $|A \cap B| \leqslant d-1$.*

Proof.

(P1) Fix $\mathbf{i} \in [n]_{\leqslant d}$ and let $s = |\mathbf{i}|$. There are $\binom{n-s}{d-s}$ d-sets in $[n]_d$ containing $\mathbf{i}$. Every tribe $A \in \mathcal{F}_n$ such that $\mathbf{i} \subseteq A$ covers exactly $\binom{k-s}{d-s}$ d-sets containing $\mathbf{i}$. Therefore, by $(**)$,

$$\left|\{A \in \mathcal{F}_n \colon \mathbf{i} \subseteq A\}\right| \leqslant \frac{\binom{n-s}{d-s}}{\binom{k-s}{d-s}} \leqslant \frac{n^{d-s}}{k^{d-s}} \leqslant \left(\frac{n}{\log_2 n}\right)^{d-s}.$$

(P2) The upper bound follows from (P1) applied to $\mathbf{i} = \varnothing$. For the lower bound note that every tribe covers $\binom{k}{d}$ d-sets. Hence, by $(*)$,

$$t \geqslant \frac{\frac{1}{2}\binom{n}{d}}{\binom{k}{d}} \geqslant \frac{n^d}{2 \cdot k^d} \geqslant \frac{1}{2d^d}\left(\frac{n}{\log_2 n}\right)^d = \frac{1}{2d^d} \cdot 2^k,$$

and the statement follows with $c_1 = \frac{1}{2d^d}$.

(P3) This follows from $(**)$. □

PDF Help

We now turn to proving specific bounds on H_n, which will give us the desired sharpness. For that let us introduce some convenient notation for speaking about the inputs of H_n.

DEFINITION 5.4.

For $A \subseteq [n]$ we define the event $\mathcal{T}(A) \subseteq \{-1,1\}^n$ by

$$\mathcal{T}(A) := \{x \in \{-1,1\}^n\colon \ for\ every\ i \in A\ there\ is\ x_i = 1\}.$$

Clearly, $H_n(x) = 1$ if and only if $x \in \bigcup_{A \in \mathcal{F}_n} \mathcal{T}(A)$.

PROPOSITION 5.5.

There is a constant c_2 depending only on d such that for every nonempty $\mathbf{i} \subseteq [n]$ of size at most d,

$$\mathrm{JInf}_{\mathbf{i}}(H_n) \leqslant c_2 \left(\frac{\log_2 n}{n}\right)^{|\mathbf{i}|}.$$

Proof.

Consider $x \in \{-1,1\}^n$ such that every $i \in \mathbf{i}$ is $\mathbf{i}$-pivotal for H_n on x. By the definition of the $\mathbf{i}$-pivotality, for every i there is $y \in \{-1,1\}^n$ such that $y_j = x_j$ for every $j \in [n] \setminus \mathbf{i}$ and

$$H_n(y^{i \to +1}) \neq H_n(y^{i \to -1}).$$

As H_n is increasing, the above may be rewritten as

$$H_n(y^{i \to +1}) = 1 \quad \text{and} \quad H_n(y^{i \to -1}) = -1.$$

PDF
Help

It follows that there is a tribe $A \in \mathcal{F}_n$ (which may not be unique) such that $i \in A$ and $x \in \mathcal{T}(A \setminus \mathbf{i})$.

Let $B_{\mathbf{i}}$ denote the set of all partitions of $\mathbf{i}$. Taking into account that the above property holds for all $i \in \mathbf{i}$, we conclude that there is a partition $S \in B_{\mathbf{i}}$ and distinct tribes $A_{\mathbf{j}}$ for $\mathbf{j} \in B_{\mathbf{i}}$, so that $x \in \mathcal{T}(A_{\mathbf{j}} \setminus \mathbf{i})$. Hence,

$$\begin{aligned} \mathrm{JInf}_{\mathbf{i}}(H_n) &\leqslant \mathbb{P}\left(\bigcup_{S \in B_{\mathbf{i}}} \bigcup_{\substack{\{A_{\mathbf{j}}\}_{\mathbf{j} \in S} \subseteq \mathcal{F}_n : \\ \forall \mathbf{j} \in S : \mathbf{j} \subseteq A_{\mathbf{j}}}} \bigcap_{\mathbf{j} \in S} \mathcal{T}(A_{\mathbf{j}} \setminus \mathbf{i}) \right) \\ &\leqslant \sum_{S \in B_{\mathbf{i}}} \sum_{\substack{\{A_{\mathbf{j}}\}_{\mathbf{j} \in S} \subseteq \mathcal{F}_n : \\ \forall \mathbf{j} \in S : \mathbf{j} \subseteq A_{\mathbf{j}}}} \mathbb{P}\left(\bigcap_{\mathbf{j} \in S} \mathcal{T}(A_{\mathbf{j}} \setminus \mathbf{i}) \right). \end{aligned} \tag{9}$$

As the number of partitions of $\mathbf{i}$ is bounded by some function of d, by (9) it is enough to show that for every partition $S \in B_{\mathbf{i}}$ we have

$$\sum_{\substack{\{A_{\mathbf{j}}\}_{\mathbf{j} \in S} \subseteq \mathcal{F}_n : \\ \forall \mathbf{j} \in S : \mathbf{j} \subseteq A_{\mathbf{j}}}} \mathbb{P}\left(\bigcap_{\mathbf{j} \in S} \mathcal{T}(A_{\mathbf{j}} \setminus \mathbf{i}) \right) \leqslant \widetilde{c_2} \cdot \left(\frac{\log_2 n}{n} \right)^r. \tag{10}$$

Clearly, $\bigcap_{\mathbf{j} \in S} \mathcal{T}(A_{\mathbf{j}} \setminus \mathbf{i}) = \mathcal{T}\left(\bigcup_{\mathbf{j} \in S} A_{\mathbf{j}} \setminus \mathbf{i} \right)$. Further, as $|S| \leqslant |\mathbf{i}| \leqslant d$ and (P3), by the inclusion-exclusion principle,

$$\left| \bigcup_{\mathbf{j} \in S} A_{\mathbf{j}} \setminus \mathbf{i} \right| \geqslant |S| k - \binom{|S|}{2} (d-1) - d \geqslant |S| k - d^3,$$

so that

PDF
Help

$$\mathbb{P}\left(\bigcap_{\mathbf{j}\in S}\mathcal{T}(A_{\mathbf{j}}\setminus\mathbf{i})\right)\leqslant 2^{-|S|k+d^3}=2^{d^3}\cdot\left(\frac{\log_2 n}{n}\right)^{d|S|}. \tag{11}$$

We now prove that the number of terms in the sum in (10) is sufficiently small. Indeed, because of (P1), for every $\mathbf{j}\in S$,

$$\left|\{A_{\mathbf{j}}\in\mathcal{F}_n\colon \mathbf{j}\subseteq A_{\mathbf{j}}\}\right|\leqslant\left(\frac{n}{\log_2 n}\right)^{d-|\mathbf{j}|},$$

so that

$$\left|\{\{A_{\mathbf{j}}\}_{\mathbf{j}\in S}\subseteq\mathcal{F}_n\colon \forall\,\mathbf{j}\in S\text{ there is }\mathbf{j}\subseteq A_{\mathbf{j}}\}\right|\leqslant\prod_{\mathbf{j}\in S}\left(\frac{n}{\log_2 n}\right)^{d-|\mathbf{j}|}=\left(\frac{n}{\log_2 n}\right)^{d|S|} \tag{12}$$

Finally, by (11) and (12),

$$\sum_{\substack{\{A_{\mathbf{j}}\}_{\mathbf{j}\in S}\subseteq\mathcal{F}_n\colon\\ \forall\mathbf{j}\in S\colon\mathbf{j}\subseteq A_{\mathbf{j}}}}\mathbb{P}\left(\bigcap_{\mathbf{j}\in S}\mathcal{T}(A_{\mathbf{j}}\setminus\mathbf{i})\right)\leqslant\left(\frac{n}{\log_2 n}\right)^{d|S|-r}\cdot 2^{d^3}\left(\frac{\log_2 n}{n}\right)^{d|S|}=2^{d^3}\left(\frac{\log_2 n}{n}\right)^{r},$$

which yields the statement with $c_2=|B_{\mathbf{i}}|\cdot 2^{d^3}\leqslant d^d 2^{d^3}$. □

In the next two propositions, we prove that H_n attains both values ± 1 with at least positi constant probability. For the proof of Proposition 5.8 we use the Harris inequality for decrea which we recall below. In Proposition 5.9 the argument proceeds via a standard second moment method. Afterwards, we finally prove Theorem 5.10.

Definition 5.6.

We say that an event $S \subseteq \{-1,1\}^n$ is decreasing *if for every $x, y \in \{-1,1\}^n$ the conditions $x \in S$ and $x \geqslant y$ coordinatewise imply that $y \in S$.*

THEOREM 5.7 (see [7]).

If $S_1, S_2, \ldots, S_m \subseteq \{-1,1\}^n$ are all decreasing events, then

$$\mathbb{P}\left(\bigcap_{i\in[m]} S_i\right) \geqslant \prod_{i\in[m]} \mathbb{P}(S_i).$$

PROPOSITION 5.8.

There is a universal constant $c_3 > 0$ such that $\mathbb{P}(H_n = -1) \geqslant c_3$.

Proof.

Let $\mathcal{F}_n = \{A_i : i \in [t]\}$. Denote by X^c the complement of the event $X \subseteq \{-1,1\}^n$.

We have $\{H_n = -1\} = \bigcap_{i\in[t]} \mathcal{T}(A_i)^c$. Further, for every $i \in [t]$, the event $\mathcal{T}(A_i)^c$ is decreasing. Hence, by the Harris inequality for the events $\mathcal{T}(A_1)^c, \ldots, \mathcal{T}(A_t)^c$,

$$\mathbb{P}(H_n = -1) = \mathbb{P}\left(\bigcap_{i\in[t]} \mathcal{T}(A_i)^c\right) \geqslant \prod_{i\in[t]} \mathbb{P}\Big(\mathcal{T}(A_i)^c\Big) = (1-2^{-k})^t \geqslant \mathrm{e}^{-t2^{-k+1}} \overset{(\mathrm{P2})}{\geqslant} \mathrm{e}^{-2},$$

where we used the bound $1 - x \geqslant \mathrm{e}^{-2x}$ which is valid for $x \in [0, \frac{1}{4}]$. □

PROPOSITION 5.9.

There is a constant $c_4 > 0$ depending only on d such that $\mathbb{P}(H_n = +1) \geqslant c_4$.

Proof.

Let $\mathcal{F}_n = \{A_i : i \in [t]\}$. Further, let $N = \sum_{i\in[t]} N_i$, where $N_i = \mathbb{1}_{\mathcal{T}(A_i)}$ are random variables.

PDF Help

Observe that $\{H_n = +1\} = \{N > 0\}$. By the Paley–Zygmund inequality,

$$\mathbb{P}(H_n = +1) = \mathbb{P}(N > 0) \geqslant \frac{(\mathbb{E}N)^2}{\mathbb{E}[N^2]} = \frac{(t\mathbb{E}N_1)^2}{\sum_{1\leqslant i\leqslant t}\mathbb{E}[N_i^2] + 2\sum_{1\leqslant i<j\leqslant t}\mathbb{E}[N_iN_j]}. \tag{13}$$

Clearly, $\mathbb{E}N_1 = \mathbb{E}[N_i^2] = \mathbb{P}(\mathcal{T}(A_i)) = 2^{-k}$. For the expectations of the products, note that

$$\mathbb{E}[N_iN_j] = \mathbb{P}(\mathcal{T}(A_i)\cap\mathcal{T}(A_j)) = \mathbb{P}(\mathcal{T}(A_i\cup A_j)).$$

By (P3), $|A_i \cup A_j| > 2k - d$ for distinct $i, j \in [t]$, so $\mathbb{E}[N_iN_j] \leqslant 2^{-2k+d}$. Hence, (13) yields

$$\mathbb{P}(H_n = +1) \geqslant \frac{t^2 2^{-2k}}{t2^{-k} + t(t-1)2^{-2k+d}} \geqslant \frac{t2^{-k}}{1 + 2^d\cdot t2^{-k}} \overset{\text{(P2)}}{\geqslant} \frac{c_1}{1+2^dc_1},$$

and the statement follows with $c_4 = \frac{c_1}{1+2^dc_1}$. □

THEOREM 5.10.

For every $d \geqslant 2$ there is a constant c_5 depending only on d such that

$$\mathrm{JInf}_{\mathbf{i}}(H_n) \leqslant c_5\mathbf{W}^{\geqslant d}(H_n)\left(\frac{\log_2 n}{n}\right)^d$$

for all $\mathbf{i} \in [n]^d$. In particular, for every $d \geqslant 2$ Theorem 1.5 is asymptotically sharp.

Proof.

PDF Help

By Proposition 5.5 we have $\mathbf{JInf}_{\mathbf{i}}(H_n) \leqslant c_2\left(\frac{\log_2 n}{n}\right)^d$ for all $\mathbf{i} \in [n]^d$. Hence, it is enough to show that there is a constant $c_6 > 0$ such that for sufficiently large n we have $\mathbf{W}^{\geqslant d}(H_n) > c_6$.

Since H_n is Boolean,

$$\mathbf{W}^{\geqslant d}(H_n) = 1 - \widehat{H^n}(\varnothing)^2 - \sum_{r=1}^{d-1} \mathbf{W}^{=r}(H_n). \tag{14}$$

By Propositions 5.8 and 5.9,

$$1 - \widehat{H^n}(\varnothing)^2 = 4\mathbb{P}(H_n = +1)\mathbb{P}(H_n = -1) \geqslant 4c_3c_4. \tag{15}$$

Further, for every $r \in [d-1]$, by Proposition 3.4,

$$\mathbf{W}^{=r}(H_n) = \sum_{\mathbf{i}\in[n]_r} \widehat{H^n}(\mathbf{i})^2 \leqslant 2^{2d} \sum_{\mathbf{i}\in[n]_r} \left(\mathrm{Inf}_{\mathbf{i}}(H_n)\right)^2 \leqslant 2^{2d} n^r \left(\mathrm{MaxInf}_r(H_n)\right)^2$$

and by Propositions 3.5 and 5.5 this may be further bounded above by

$$c_2^2 2^{2d} n^r \cdot \left(\frac{\log_2 n}{n}\right)^{2r} = c_2^2 2^{2d} \cdot \frac{(\log_2 n)^{2r}}{n^r},$$

so $\sum_{r=1}^{d-1} \mathbf{W}^{=r}(H_n) < 2c_3c_4$ for sufficiently large n. Hence, by (14) and (15), for sufficiently large n, we have $\mathbf{W}^{\geqslant d}(H_n) > c_3c_4$, as desired. □

6. The FKN-type theorem for the multibit influences.

We now turn our attention to Theorem 1.7. Roughly speaking, it says that a Boolean function which is close to a d-degree function in the sense of having small $(d+1)$-degree influences is necessarily

close to a d-degree Boolean function. Such a statement might be seen as a variation of the FKN Theorem [6], or its higher order generalization—the Kindler–Safra Theorem [10]—Theorem 1.4. Indeed, the Kindler–Safra Theorem ensures that a Boolean function which is close to a d-degree function in the sense of having small $(d+1)$-order variance $\mathbf{W}^{\geqslant d+1}$ is necessarily close to a d-degree Boolean function. As it turns out, the assumption of having small $(d+1)$-degree influences is strictly stronger than having small $(d+1)$-order variance $\mathbf{W}^{\geqslant d+1}$. However, this stronger assumption yields a better approximation than the one provided by the Kindler–Safra Theorem.

Additionally, Theorem 1.7 is a direct generalization of Oleszkiewicz's Theorem 1.2, which can be recovered by taking $d=1$. In his paper, [13], Oleszkiewicz presents two separate proofs of Theorem 1.2; one solely uses hypercontraction, the other the log-Sobolev inequality. Our proof of Theorem 1.7 combines the two approaches. It seems unlikely to us that just using one of them would be sufficient in this higher order case.

In the following section we lay the foundation for the proof of Theorem 1.7. First, in Proposition 6.1 we prove an integral representation of $\mathrm{Inf}_{\mathbf{j}}(g)-\hat{g}(\mathbf{j})^2$. In Propositions 6.2 and 6.5 we prove some approximation results regarding influences and Fourier–Walsh coefficients. Further, in Proposition 6.6 we prove that the Fourier–Walsh coefficients of a d-degree Boolean function lie in the $\mathbb{Z}/2^{d-1}$ lattice.

6.1. Hypercontractive propositions.

In the following proposition we prove a higher order analogue of Lemma 3.3 from [13].

PROPOSITION 6.1.

Let $n>k\geqslant 0$ be integers. For every $g\colon\{-1,1\}^n\to\mathbb{R}$ and $\mathbf{j}\in[n]_k$,

$$\mathrm{Inf}_{\mathbf{j}}(g)-\hat{g}(\mathbf{j})^2=2\sum_{\substack{\mathbf{i}\in[n]_{k+1}:\\ \mathbf{j}\subseteq\mathbf{i}}}\int_0^\infty e^{-2t}\|P_t\partial_{\mathbf{i}}g\|^2dt.$$

Proof.

Let $\mathbf{j}\in[n]_k$. Observe that $\widehat{(\partial_{\mathbf{j}}g)}(\varnothing)=\mathbb{E}[\partial_{\mathbf{j}}g]=\hat{g}(\mathbf{j})$. Thus, by the definition of influence,

PDF Help

$$\mathrm{Inf}_{\mathbf{j}}(g) - \hat{g}(\mathbf{j})^2 = \|\partial_{\mathbf{j}} g\|_2^2 - \widehat{(\partial_{\mathbf{j}} g)}(\varnothing)^2 = \mathbf{W}^{\geqslant 1}(\partial_{\mathbf{j}} g).$$

Further, Proposition 4.1 applied to the function $f = \partial_{\mathbf{j}} g$ and $d = 1$ yields

$$\mathbf{W}^{\geqslant 1}(\partial_{\mathbf{j}} g) = 2 \sum_{i \in [n]} \int_0^\infty \mathrm{e}^{-2t} \|P_t \partial_i (\partial_{\mathbf{j}} g)\|_2^2 \mathrm{d}t = 2 \sum_{\substack{\mathbf{i} \in [n]_{k+1}: \\ \mathbf{j} \subseteq \mathbf{i}}} \int_0^\infty \mathrm{e}^{-2t} \|P_t \partial_{\mathbf{i}} g\|^2 \mathrm{d}t,$$

which finishes the proof. □

The next proposition relates how close is a Fourier–Walsh coefficient to the respective influence in terms of influences of one degree higher. We prove it via Proposition 4.2 which uses hypercontraction.

PROPOSITION 6.2.

Let n, d, k be integers such that $n \geqslant 2$, $d \geqslant 1$, and $n > k \geqslant 0$. Assume that $g\colon \{-1,1\}^n \to \mathbb{R}$ and $\beta \in [0,1]$ satisfy

$$\mathrm{MaxInf}_{k+1}(g) \leqslant \beta \left(\frac{\ln n}{n}\right)^{k+1} \quad \textit{and for every } \mathbf{i} \in [n]_{k+1} \textit{ there is } \mathrm{Im}(\partial_{\mathbf{i}} g) \subseteq \mathbb{Z}/2^{d-}$$

Then for every $\mathbf{j} \in [n]_k$,

$$\mathrm{Inf}_{\mathbf{j}}(g) - \hat{g}(\mathbf{j})^2 \leqslant 2^{2d} \beta \left(\frac{\ln n}{n}\right)^{k}.$$

Proof.

PDF Help

Let $\mathbf{i} \in [n]_{k+1}$ be arbitrary. Then we have $\operatorname{Inf}_{\mathbf{i}}(g) \leqslant \operatorname{MaxInf}_{k+1}(g) \leqslant 1$. Therefore, applying Proposition 4.2 to the function $f = g, l = 1$ and $\mathbf{i}$ we get

$$\int_0^\infty \mathrm{e}^{-2t} \|P_t \partial_{\mathbf{i}} g\|^2 \mathrm{d}t \leqslant \frac{2^{2d-2} \operatorname{Inf}_{\mathbf{i}}(g)}{\ln(1/\operatorname{Inf}_{\mathbf{i}}(g))}. \tag{16}$$

Fix $\mathbf{j} \in [n]_k$. By Proposition 6.1, (16), and the fact that the function $t \mapsto t/\ln(1/t)$ is increasing on [0,1], we obtain

$$\begin{aligned}\operatorname{Inf}_{\mathbf{j}}(g) - \hat{g}(\mathbf{j})^2 &= 2 \sum_{\substack{\mathbf{i} \in [n]_{k+1}: \\ \mathbf{j} \subseteq \mathbf{i}}} \int_0^\infty \mathrm{e}^{-2t} \|P_t \partial_{\mathbf{i}} g\|^2 \mathrm{d}t \leqslant 2^{2d-1} \sum_{\substack{\mathbf{i} \in [n]_{k+1}: \\ \mathbf{j} \subseteq \mathbf{i}}} \frac{\operatorname{Inf}_{\mathbf{i}}(g)}{\ln(1/\operatorname{Inf}_{\mathbf{i}}(g))} \\ &\leqslant 2^{2d-1} \frac{n \cdot \operatorname{MaxInf}_{k+1}(g)}{\ln(1/\operatorname{MaxInf}_{k+1}(g))} \leqslant 2^{2d-1} \cdot \beta \left(\frac{\ln n}{n}\right)^k \cdot \frac{\ln n}{\ln\left(\frac{1}{\beta} \cdot \left(\frac{n}{\ln n}\right)^{k+1}\right)}\end{aligned}$$

Observe that $n \geqslant 2$, so that $\sqrt{n} \geqslant \ln n$. Thus, since $\beta \leqslant 1$, we get

$$2^{2d-1} \beta \left(\frac{\ln n}{n}\right)^k \cdot \frac{\ln n}{\ln\left(\frac{1}{\beta} \cdot \left(\frac{n}{\ln n}\right)^{k+1}\right)} \leqslant 2^{2d-1} \beta \left(\frac{\ln n}{n}\right)^k \cdot \frac{\ln n}{\ln \sqrt{n}} = 2^{2d} \beta \left(\frac{\ln n}{n}\right)^k$$

□

6.2. Log-Sobolev propositions.

In the following propositions we prove that a function that is close to a d-degree Boolean f in some specific sense, has its Fourier–Walsh coefficients close to integer multiples of $1/2^{d-1}$. Proposition 6.4 is a minor modification of Corollary 5.4 from [13].

Let us introduce some convenient notation.

PDF Help

Definition 6.3.

Let $d \geqslant 1$ be an integer. We define $V_d = [-2^{2d}, 2^{2d}] \cap (\mathbb{Z}/2^{d-1})$.

Proposition 6.4.

Let $d \geqslant 2$ be an integer. Assume $g\colon \{-1,1\}^n \to V_d$ satisfies $\mathrm{Inf}(g) < \frac{1}{2^{5d}}$. Then there exists $\eta \in \mathrm{Im}(g) \subseteq V_d$ such that

$$\mathbb{P}(g \neq \eta) \leqslant \frac{2^{2d-1}\,\mathrm{Inf}(g)}{\ln\left(\frac{1}{2^{5d}\,\mathrm{Inf}(g)}\right)}.$$

Proof.

If $\mathrm{Inf}(g) = 0$, then g is constant and in this case $\eta = g$ satisfies the statement.

Assume, henceforth, that $\mathrm{Inf}(g) > 0$. Note that

$$\sum_{\theta \in V_d} \mathbb{P}(g \neq \theta) = \sum_{\theta \in V_d} (1 - \mathbb{P}(g = \theta)) = |V_d| - 1 = 2^{3d}.$$

Therefore, we may pick $\eta \in V_d$ such that

$$\mathbb{P}(g \neq \eta) \leqslant \frac{|V_d| - 1}{|V_d|}. \tag{17}$$

Let $h = \mathbb{1}_{\{g \neq \eta\}}$. We claim that

$$\mathrm{Inf}(h) \leqslant 2^{2d-2} \cdot \mathrm{Inf}(g). \tag{18}$$

PDF Help

Indeed, note that

$$|\partial_i h(x)| = \frac{1}{2}\left|h(x^{i\to+1}) - h(x^{i\to-1})\right| = \frac{1}{2}\left|\mathbb{1}_{V_d\setminus\{\eta\}}(g(x^{i\to+1})) - \mathbb{1}_{V_d\setminus\{\eta\}}(g(x^{i\to-1}))\right|$$
$$\leqslant \frac{2^{d-1}}{2}\cdot\left|g(x^{i\to+1}) - g(x^{i\to-1})\right| = 2^{d-1}\cdot|\partial_i g(x)|,$$

where we used the fact that $\min\{|v-w|\colon v,w\in V_d, v\neq w\} = 1/2^{d-1}$. This inequality yields that for every $i\in[n]$ it holds that $\|\partial_i h\|_2^2 \leqslant 2^{2d-2}\cdot\|\partial_i g\|_2^2$. The claim follows from summing this relation over $i\in[n]$.

Observe that $\mathbb{E}h = \mathbb{P}(g\neq\eta)$. By claim (18), the log-Sobolev Corollary 3.10 applied to the $\{0,1\}$-valued function h, the monotonicity of $\ln$ with (17), and the bound $\ln(1+x)\geqslant\frac{x}{2}$ for $x\in(0,1)$, we get

$$2^{2d-2}\operatorname{Inf}(g) \geqslant \operatorname{Inf}(h) \geqslant \frac{1}{2}\mathbb{E}h\cdot\ln\left(\frac{1}{\mathbb{E}h}\right) \geqslant \frac{1}{2}\mathbb{E}h\cdot\ln\left(1+\frac{1}{|V_d|-1}\right) \geqslant \frac{\mathbb{E}h}{2^{3d+2}}. \tag{19}$$

Inequality (19) implies that $\mathbb{E}h\leqslant 2^{5d}\operatorname{Inf}(g)$. Using this, we may improve the lower bound of $\ln(1/\mathbb{E}h)$ in (19) as follows:

$$2^{2d-2}\operatorname{Inf}(g) \geqslant \frac{1}{2}\mathbb{E}h\cdot\ln\left(\frac{1}{\mathbb{E}h}\right) \geqslant \frac{1}{2}\mathbb{E}h\cdot\ln\left(\frac{1}{2^{5d}\operatorname{Inf}(g)}\right). \tag{20}$$

The restriction $\operatorname{Inf}(g) < 2^{-5d}$ guarantees that $\ln(\frac{1}{2^{5d}\operatorname{Inf}(g)}) > 0$. Recalling that $\mathbb{E}h = \mathbb{P}(g\neq\eta)$ and rewriting (20) completes the proof. □

PROPOSITION 6.5.

Let n, d, k be integers such that $n > k \geqslant 1$ and $d\geqslant 2$. Assume that $g\colon\{-1,1\}^n\to\mathbb{R}$ and $\beta\in[0,\frac{1}{2^{6d}})$ satisfy

$$\mathrm{MaxInf}_{k+1}(g) \leqslant \beta \left(\frac{\ln n}{n}\right)^{k+1} \quad \textit{and for every } \mathbf{j} \in [n]_k \textit{ we have } \mathrm{Im}(\partial_{\mathbf{j}} g) \subseteq V_d.$$

Then for every $\mathbf{j} \in [n]_k$ *there exists* $\eta_{\mathbf{j}} \in \mathrm{Im}(\partial_{\mathbf{j}} g) \subseteq V_d$ *such that*

$$|\hat{g}(\mathbf{j}) - \eta_{\mathbf{j}}| \leqslant 2^{4d+2} \beta \left(\frac{\ln n}{n}\right)^{k}.$$

Proof.

Fix $\mathbf{j} \in [n]_k$. We will apply Proposition 6.4 to the function $\partial_{\mathbf{j}} g \colon \{-1,1\}^n \to V_d$. Let us check the assumption that $\mathrm{Inf}(\partial_{\mathbf{j}} g) < \frac{1}{2^{5d}}$. Indeed, we have $\sqrt{n} \geqslant \ln n$ and $\beta < \frac{1}{2^{5d}}$, so

$$\begin{aligned} \mathrm{Inf}(\partial_{\mathbf{j}} g) = \sum_{i \in [n]} \mathrm{Inf}_i(\partial_{\mathbf{j}} g) = \sum_{\substack{\mathbf{i} \in [n]_{k+1} \\ \mathbf{j} \subseteq \mathbf{i}}} \mathrm{Inf}_{\mathbf{i}}(g) \leqslant \sum_{\substack{\mathbf{i} \in [n]_{k+1} \\ \mathbf{j} \subseteq \mathbf{i}}} \beta \left(\frac{\ln n}{n}\right)^{k+1} \\ \leqslant \beta \frac{(\ln n)^{k+1}}{n^k} < \frac{1}{2^{5d}}. \end{aligned} \tag{21}$$

Therefore, we get $\eta_{\mathbf{j}} \in \mathrm{Im}(\partial_{\mathbf{j}} g) \subseteq V_d$ satisfying

$$\mathbb{P}(\partial_{\mathbf{j}} g \neq \eta_{\mathbf{j}}) \leqslant \frac{2^{2d-1} \mathrm{Inf}(\partial_{\mathbf{j}} g)}{\ln\left(\frac{1}{2^{5d} \mathrm{Inf}(\partial_{\mathbf{j}} g)}\right)} \leqslant 2^{2d-1} \beta \left(\frac{\ln n}{n}\right)^{k} \cdot \frac{\ln n}{\ln\left(\frac{n^k}{(\ln n)^{k+1}}\right)} \leqslant 2^{2d+1} \beta \left(\frac{\ln} \tag{22}$$

where we used an intermediate bound from (21) and the inequality $2^{5d}\beta < 1$.

Now observe that

$$\begin{aligned}\hat{g}(\mathbf{j}) = \mathbb{E}[\partial_{\mathbf{j}} g] &= \eta_{\mathbf{j}} \mathbb{P}(\partial_{\mathbf{j}} g = \eta_{\mathbf{j}}) + \mathbb{E}[(\partial_{\mathbf{j}} g) \mathbb{1}_{\{\partial_{\mathbf{j}} g \neq \eta_{\mathbf{j}}\}}] \\ &= \eta_{\mathbf{j}} (1 - \mathbb{P}(\partial_{\mathbf{j}} g \neq \eta_{\mathbf{j}})) + \mathbb{E}[(\partial_{\mathbf{j}} g) \mathbb{1}_{\{\partial_{\mathbf{j}} g \neq \eta_{\mathbf{j}}\}}].\end{aligned}$$

Thus, as $|\eta_{\mathbf{j}}| \leqslant 2^{2d}$ and $\|\partial_{\mathbf{j}} g\|_\infty := \max_{x \in \{-1,1\}^n} |\partial_{\mathbf{j}}(x)| \leqslant 2^{2d}$, we get

$$|\hat{g}(\mathbf{j}) - \eta_{\mathbf{j}}| \leqslant (|\eta_{\mathbf{j}}| + \|\partial_{\mathbf{j}} g\|_\infty) \mathbb{P}(\partial_{\mathbf{j}} g \neq \eta_{\mathbf{j}}) \leqslant 2^{4d+2} \beta \left(\frac{\ln n}{n} \right)^k.$$

□

6.3. Coefficients of a d-degree function.

In this section we prove that the Fourier–Walsh coefficients of a d-degree Boolean function lie in the $\mathbb{Z}/2^{d-1}$ lattice. This result in print has appeared in [9, Proposition 3.1], but we include our rather short proof to keep the work self-contained.

PROPOSITION 6.6.

Let $d \geqslant 1$ be an integer, and let $g\colon \{-1,1\}^n \to \{-1,1\}$ be a d-degree function. Then for every $\mathbf{i} \subseteq [n]$ we have $\hat{g}(\mathbf{i}) \in \mathbb{Z}/2^{d-1}$. Consequently, at most 2^{2d-2} of the coefficients $\hat{g}(\mathbf{i})$ are nonzero.

Proof.

We have $\hat{g}(\mathbf{i}) = 0$ for $\mathbf{i} \in [n]_{>d}$, so it is enough to resolve the $\mathbf{i} \in [n]_{\leqslant d}$ case.

Assume, for a contradiction, that there is $\mathbf{i} \in [n]_{\leqslant d}$ with $\hat{g}(\mathbf{i}) \notin \mathbb{Z}/2^{d-1}$, and choose such a set $\mathbf{i}$ of maximal cardinality. We have

$$\partial_{\mathbf{i}} g = \hat{g}(\mathbf{i}) + \sum_{\substack{\mathbf{j} \in [n]_{\leqslant d} \\ \mathbf{i} \subsetneq \mathbf{j}}} \hat{g}(\mathbf{j}) \chi_{\mathbf{j} \setminus \mathbf{i}}. \tag{23}$$

PDF Help

Since g is Boolean, by Corollary 3.3, $\mathrm{Im}(\partial_{\mathbf{i}} g) \subseteq \mathbb{Z}/2^{|\mathbf{i}|-1} \subseteq \mathbb{Z}/2^{d-1}$. Further, by the choice of $\mathbf{i}$, we have $\hat{g}(\mathbf{j}) \in \mathbb{Z}/2^{d-1}$ for all $\mathbf{j} \in [n]_{\leqslant d}$ such that $\mathbf{j} \supsetneq \mathbf{i}$. Hence, all of the terms but one in (23) are integer multiples of $1/2^{d-1}$, which yields $\hat{g}(\mathbf{i}) \in \mathbb{Z}/2^{d-1}$, a contradiction.

As by Parseval's Theorem $\sum_{\mathbf{i} \subseteq [n]} \hat{g}(\mathbf{i})^2 = 1$, clearly at most 2^{2d-2} of the coefficients $\hat{g}(\mathbf{i})$ are nonzero. □

7. Proof of Theorem 1.7.

We are now ready to prove Theorem 1.7. Let us sketch the plan of the proof.

First, thanks to Theorem 1.5 and the Kindler–Safra Theorem 1.4, the assumption (24) of small $(d+1)$-degree influences gives a candidate d-degree Boolean function g approximating f. All we need to prove is that this function g is sufficiently close to f in the sense of the bounds (25). We show this by a strengthened induction, where in parallel to proving the bound we construct a scheme of (quasi-Boolean) functions with small influences which have good approximating properties owing Propositions 6.2 and 6.5.

We will prove a slightly (and equivalently) rephrased statement of Theorem 1.7.

THEOREM 7.1.

Let $n > d \geqslant 1$. There are positive constants C_3, C_4 depending only on d such that if $f\colon \{-1,1\}^n \to \{-1,1\}$ satisfies

$$\mathrm{MaxInf}_{d+1}(f) \leqslant \alpha \left(\frac{\ln n}{n}\right)^{d+1} \tag{24}$$

for some $\alpha \in (0, C_3)$, then there exists a d-degree function $g\colon \{-1,1\}^n \to \{-1,1\}$ such that for every $\mathbf{j} \in [n]_{\leqslant d}$ we have

$$|\hat{f}(\mathbf{j}) - \hat{g}(\mathbf{j})| \leqslant C_4 \alpha \left(\frac{\ln n}{n}\right)^{|\mathbf{j}|}. \tag{25}$$

PDF Help

Proof.

Observe that the bound (24) and Theorem 1.5 yield $\mathbf{W}^{\geqslant d+1}(f) \leqslant 10\alpha$. Hence, the Kindler–Safra Theorem (Theorem 1.4) implies the existence of a d-degree $g\colon \{-1,1\}^n \to \{-1,1\}$ such that

$$\|f-g\|_2^2 \leqslant 10 C_{KS}\alpha. \tag{26}$$

We will prove that the g obtained satisfies the bound (25) for $C_3 = \min\{2^{-10d^2}, (2^{2d}10C_{KS})^{-1}\}$ and $C_4 = \max\{2^{8d^2}, 10C_{KS}\}$.

First, let us note that the bound (25) holds for $\mathbf{j} = \varnothing$. Indeed,

$$|\hat{f}(\varnothing) - \hat{g}(\varnothing)| = |\mathbb{E}[f-g]| \leqslant \mathbb{E}[|f-g|] \leqslant \mathbb{E}[(f-g)^2] = \|f-g\|_2^2 \leqslant 10C_{KS}\alpha \leqslant \tag{27}$$

Define $f_1, \dots, f_d\colon \{-1,1\}^n \to \mathbb{R}$ by $f_d := f$ and

$$f_k \;:=\; f - \sum_{\substack{\mathbf{i}\subseteq[n]\\ k<|\mathbf{i}|\leqslant d}} \hat{g}(\mathbf{i})\chi_{\mathbf{i}} \;=\; f_{k+1} - \sum_{\mathbf{i}\in[n]_{k+1}} \hat{g}(\mathbf{i})\chi_{\mathbf{i}}. \tag{28}$$

Proposition 6.6 yields $\hat{g}(\mathbf{i}) \in \mathbb{Z}/2^{d-1}$ for every $\mathbf{i} \in [n]_{\leqslant d}$. Fix $\mathbf{j} \in [n]_{\leqslant d}$. By Corollary 3.3 we have $\operatorname{Im}(\partial_{\mathbf{j}} f) \subseteq \mathbb{Z}/2^{d-1}$. In turn, due to the Fourier–Walsh expansion of $\partial_{\mathbf{j}}$, we have $\operatorname{Im}(\partial_{\mathbf{j}}\chi_{\mathbf{i}}) \subseteq \{-1,0,1\}$. Therefore, (28) yields $\operatorname{Im}(\partial_{\mathbf{j}} f_k) \subseteq \mathbb{Z}/2^{d-1}$.

On the other hand, the second part of Proposition 6.6 implies the pointwise bound

$$|f_k| \leqslant |f| + \sum_{\substack{\mathbf{i}\subseteq[n]\\ k<|\mathbf{i}|\leqslant d}} |\hat{g}(\mathbf{i})| \leqslant 1 + 2^{2d-2} < 2^{2d}.$$

PDF Help

Hence,

$$\operatorname{Im}(\partial_{\mathbf{j}} f_k) \subseteq [-2^{2d}, 2^{2d}] \cap \mathbb{Z}/2^{d-1} = V_d \text{ for every } k \in [d] \text{ and } \mathbf{j} \in [n]_{\leqslant d}. \tag{29}$$

Let $\alpha_d = \alpha$ and recursively define $\alpha_k = 2^{8d}\alpha_{k+1}$ for $k \in [d-1]$. Due to the choice of C_3, we have

$$\alpha_k = 2^{8d(d-k)}\alpha \quad \text{and} \quad \alpha = \alpha_d \leqslant \alpha_{d-1} \leqslant \cdots \leqslant \alpha_1 \leqslant \frac{1}{2^{10d}} < 1.$$

We claim that for every $k \in [d]$ the following conditions are satisfied:

$(\mathrm{C1}_k)$ $\operatorname{MaxInf}_{k+1}(f_k) \leqslant \alpha_k (\frac{\ln n}{n})^{k+1}$;

$(\mathrm{C2}_k)$ for every $\mathbf{j} \in [n]_k$ we have

$$\operatorname{Inf}_{\mathbf{j}}(f_k) - \hat{f}_k(\mathbf{j})^2 \leqslant 2^{2d}\alpha_k (\frac{\ln n}{n})^k;$$

$(\mathrm{C3}_k)$ for every $\mathbf{j} \in [n]_k$ we have

$$|\hat{f}_k(\mathbf{j}) - \hat{g}(\mathbf{j})| = |\hat{f}(\mathbf{j}) - \hat{g}(\mathbf{j})| \leqslant 2^{4d+2}\alpha_k \left(\frac{\ln n}{n}\right)^k.$$

Observe that $2^{4d+2}\alpha_k \leqslant 2^{8d^2}\alpha$, so that (C3) and (27) imply the theorem with the chosen constant C_4.

In order to prove the claim, we run a downwards induction on k. We perform it in two steps. First, we prove that for every $k \in [d]$ the condition ($\mathrm{C1}_k$) implies ($\mathrm{C2}_k$) and ($\mathrm{C3}_k$). Second, we prove that conditions ($\mathrm{C2}_k$) and ($\mathrm{C3}_k$) imply ($\mathrm{C1}_{k-1}$).

The base condition ($\mathrm{C1}_d$) is simply the assumption (24).

PDF Help

Let $\mathbf{j} \in [n]_k$ and assume (C1$_k$) holds.

(C2$_k$): Because of (C1$_k$) and (29), we are allowed to apply Proposition 6.2 to the function f_k and $\beta = \alpha_k$, which yields the condition.

(C3$_k$): Because of (C1$_k$) and (29), we are allowed to apply Proposition 6.5 to the function f_k and $\beta = \alpha_k$. As a result, we obtain $\eta_{\mathbf{j}} \in V_d$ such that

$$|\hat{f}_k(\mathbf{j}) - \eta_{\mathbf{j}}| \leqslant 2^{4d+2}\alpha_k \left(\frac{\ln n}{n}\right)^k. \tag{30}$$

Since $|\mathbf{j}| \leqslant k$, the definition of f_k yields $\hat{f}_k(\mathbf{j}) = \hat{f}(\mathbf{j})$. Thus, it is enough to show that $\hat{g}(\mathbf{j}) = \eta_{\mathbf{j}}$.

Let us observe that for every $\mathbf{i} \subseteq [n]$ we have

$$|\hat{f}(\mathbf{i}) - \hat{g}(\mathbf{i})| = \sqrt{(\hat{f}(\mathbf{i}) - \hat{g}(\mathbf{i}))^2} \leqslant \|f - g\|_2 \overset{(25)}{\leqslant} \sqrt{10C_{KS}\alpha} < \frac{1}{2^d}, \tag{31}$$

so by (30) and inequality $2^{4d+2}\alpha_k < \frac{1}{2^d}$,

$$|\eta_{\mathbf{j}} - \hat{g}(\mathbf{j})| \leqslant |\hat{f}(\mathbf{j}) - \hat{g}(\mathbf{j})| + |\hat{f}(\mathbf{j}) - \eta_{\mathbf{j}}| < \frac{1}{2^{d-1}}. \tag{32}$$

Recall that $\eta_{\mathbf{j}} \in V_d$ and that $\hat{g}(\mathbf{j})$ is a coefficient of a d-degree Boolean function. Hence, both $\eta_{\mathbf{j}}$ and $\hat{g}(\mathbf{j})$ are integer multiples of $\frac{1}{2^{d-1}}$. Thus, (32) implies $\hat{g}(\mathbf{j}) = \eta_{\mathbf{j}}$, and the condition follows.

Turning to the second step, let us assume that (C2$_k$) and (C3$_k$) hold for some $k \in \{2, \dots,$

(C1$_{k-1}$): Let $\mathbf{i} \in [n]_k$. By the definition of f_{k-1} and conditions (C2$_k$), (C3$_k$), we get

PDF
Help

$$\begin{aligned}\operatorname{Inf}_{\mathbf{i}}(f_{k-1}) &= \operatorname{Inf}_{\mathbf{i}}(f_k) - \hat{f}_k(\mathbf{i})^2 + (\hat{f}_k(\mathbf{i}) - \hat{g}(\mathbf{i}))^2 \\ &\leqslant 2^{2d}\alpha_k \left(\frac{\ln n}{n}\right)^k + 2^{8d+4}\alpha_k^2 \left(\frac{\ln n}{n}\right)^{2k} \leqslant \alpha_{k-1}\left(\frac{\ln n}{n}\right)^k,\end{aligned}$$

as $\alpha_k \leqslant \frac{1}{2^{10d}}$. Since $\mathbf{i} \in [n]_k$ was arbitrary, the condition follows. □

Remark 7.2.

It is possible to avoid the use of the Kindler–Safra Theorem. Our original argument showed directly that the function $\sum \eta_{\mathbf{i}}\chi_{\mathbf{i}}$ is the desired d-degree Boolean g. Such an approach, although elementary, is considerably longer.

Remark 7.3.

A nontrivial example of f satisfying the assumption (24) of small d-degree influences is a slightly modified version of the hypertribe function described in section 5. The nuance to overcome in this case is making the parameter α smaller than C_3. In order to guarantee this, we may increase the size of the tribes by a constant (depending on d). After this modification, the bounds obtained in Proposition 5.3 will still be satisfied for sufficiently large n, and following the proof of Proposition 5.5, one can see that α will be suitably small. In this case, for sufficiently large d and n, Propositions 5.5 and 3.4 yield that the approximating function g is the constant -1.

We close with the proof of Corollary 1.8.

Proof of Corollary 1.8.

We show that the function g provided by Theorem 1.7 satisfies the bound. Assume $\mathbf{i} \in [n]_{\leqslant d}$ is such that $|\hat{f}(\mathbf{i})| < \frac{1}{2^d}$. Since g is d-degree, Proposition 6.6 yields that $\hat{g}(\mathbf{i})$ is an integer multiple of $\frac{1}{2^{d-1}}$. Therefore, appealing to the above proof, (31) implies $\hat{g}(\mathbf{i}) = 0$.

Thus, for all $\mathbf{i} \in [n]_{<l}$ we have $\hat{g}(\mathbf{i}) = 0$, so by Plancherel's Theorem,

PDF Help

$$\|f-g\|_2^2 = \|f\|_2^2 - \|g\|_2^2 + 2\langle g-f, g\rangle = 2\langle g-f, g\rangle = 2\sum_{\substack{\mathbf{i}\subseteq[n]\\ l\leqslant|\mathbf{i}|\leqslant d}} \left(\hat{g}(\mathbf{i}) - \hat{f}(\mathbf{i})\right)\cdot\hat{g}(\mathbf{i}).$$

Appealing again to Proposition 6.6, we know that at most 2^{2d-2} of the coefficients $\hat{g}(\mathbf{i})$ are nonzero, so using that $|\hat{g}(\mathbf{i})| \leqslant 1$ and by Theorem 1.7,

$$\|f-g\|_2^2 \leqslant 2\sum_{\substack{\mathbf{i}\subseteq[n]\\ l\leqslant|\mathbf{i}|\leqslant d}} |\hat{g}(\mathbf{i}) - \hat{f}(\mathbf{i})|\cdot|\hat{g}(\mathbf{i})| \leqslant 2^{2d-1}\cdot C_4\alpha\left(\frac{\ln n}{n}\right)^l. \tag{33}$$

The claim follows with $C_5 = 2^{2d-1}C_4$. □

Acknowledgments.

The majority of this work was written as a master's thesis at the University of Warsaw. The author would like to thank his supervisor Krzysztof Oleszkiewicz for help and for an interesting topic to research, and to Christina Goldschmidt, Peter Keevash, and Oliver Riordan for valuable suggestions and discussions.

Footnote

1 Our notation ∂_i is consistent with [13], but inconsistent with [12, 14] where the symbol D_i is used instead.

Return to text

References

1. M. Ajtai and N. Linial, The influence of large coalitions, *Combinatorica*, 13 (1993), pp. 129–145, https://doi.org/10.1007/BF01303199.

Go to Citation | Web of Science | Google Scholar | Find Full Text

2. M. Ben-Or and N. Linial, Collective coin flipping, robust voting schemes and minima of Banzhaf values, in 26th Annual Symposium on Foundations of Computer Science, 1985, pp. 408–416.

PDF Help

Show Citations | Google Scholar | Find Full Text

3. A. Biswas and P. Sarkar, Influence of a set of variables on a Boolean function, *SIAM J. Discrete Math.*, 37 (2023),

SHOW ALL REFERENCES

Recommended Content

Influence of a Set of Variables on a Boolean Function

Aniruddha Biswas ,Palash Sarkar

Abstract +

Abstract.

The influence of a variable is an important concept in the analysis of Boolean functions. The more general notion of influence of a set of variables on a Boolean function has four separate definitions in the literature. In the present work, we introduce a new definition of influence of a set of variables which is based on the auto-correlation function and develop its basic theory. Among the new results that we obtain are generalizations of the Poincaré inequality and the edge expansion property of the influence of a single variable. Further, we obtain new characterizations of resilient and bent functions using the notion of influence. We show that the previous definition of influence due to Fischer et al. [*Proceedings of the* 43*rd Symposium on Foundations of Computer Science (FOCS* 2002*)*, Vancouver, BC, Canada, IEEE Computer Society, 2002, pp. 103–112] and Blais [*Proceedings of the* 41*st Annual ACM Symposium on Theory of Computing, STOC* 2009, M. Mitzenmacher, ed., Bethesda, MD, ACM, 2009, pp. 151–158] is half the value of the auto-correlation based influence that we introduce. Regarding the other prior notions of influence, we make a detailed study of these and show that each of these definitions does not satisfy one or more desirable properties that a notion of influence may be expected to satisfy.

Approximate Graph Coloring and the Crystal with a Hollow Shadow

Lorenzo Ciardo ,Stanislav Živný

Abstract +

Abstract.

We show that approximate graph coloring is not solved by the lift-and-project hierarchy for the combination of linear programming and linear Diophantine equations. The proof is based on combinatorial tensor theory.

PDF

Help

Removing the Mask—Reconstructing a Real-Valued Field on the Sphere from a Masked Field by Spherical Fourier Analysis

Jan Hamann ,Quoc T. Le Gia ,Ian H. Sloan ,Robert S. Womersley

Abstract +

Abstract.

The paper analyzes a spectral approach to reconstructing a scalar field on the sphere, given only information about a masked version of the field, together with precise information about the (smooth) mask. The theory is developed for a general mask and later specialized to the case of an axially symmetric mask. Numerical experiments are given for the case of an axial mask motivated by the cosmic microwave background, assuming that the underlying field is a realization of a Gaussian random field with an artificial angular power spectrum of moderate degree (). The recovery is highly satisfactory in the absence of noise and even in the presence of moderate noise.

Dimension Free Nonasymptotic Bounds on the Accuracy of High-Dimensional Laplace Approximation

Vladimir Spokoiny

Abstract +

Abstract.

This paper aims at revisiting the classical results on Laplace approximation in a modern nonasymptotic and dimension-free form. Such an extension is motivated by applications to high-dimensional statistical and optimization problems. The established results provide explicit nonasymptotic bounds on the quality of a Gaussian approximation of the posterior distribution in total variation distance in terms of the so-called *effective dimension*. This value is defined as the interplay between the information contained in the data and in the prior distribution. In contrast to prominent Bernstein–von Mises results, the impact of the prior is not negligible, and it allows one to keep the effective dimension small or moderate even if the true parameter dimension is huge. We also address the issue of using a Gaussian approximation with inexact parameters, with the focus on replacing the maximum a posteriori (MAP) value by the posterior mean and designing the algorithm of Bayesian optimization based on Laplace iterations. The results are specified to the case of nonlinear regression.

A Universal Median Quasi-Monte Carlo Integration

Takashi Goda ,Kosuke Suzuki ,Makoto Matsumoto

Abstract +

Abstract.

We study quasi-Monte Carlo (QMC) integration over the multidimensional unit cube in several weighted function spaces with different smoothness classes. We consider approximating the integral of a function by the median of several integral estimates under independent and random choices of the underlying QMC point sets (either linearly scrambled digital nets or infinite-precision polynomial lattice point sets). Even though our approach does not require any information on the smoothness and weights of a target function space as an input, we can prove a probabilistic upper bound on the worst-case error for the respective weighted function space, where the failure probability converges to 0 exponentially fast as the number of [illegible] increases. Our obtained rates of convergence are nearly optimal for function spaces with finite smoothness, and we [illegible] dimension-independent super-polynomial convergence for a class of infinitely differentiable functions. This implies th[illegible] median-based QMC rule is *universal* in the sense that it does not need to be adjusted to the smoothness and the we[illegible] function spaces and yet exhibits the nearly optimal rate of convergence. Numerical experiments support our theoretical results.

PDF

Help

View full text | Download PDF

Previous Article

Next Article

Browse

Journals
E-books
Bookstore
Proceedings

Alerts

Sign up/Manage Email Alerts

Information

For Journal Authors
For Book Authors
For Librarians
Help
Terms of Use & Privacy Policy
Accessibility Statement

About

SIAM
Join SIAM
Donate to SIAM

PDF
Help

Society for Industrial and Applied Mathematics
3600 Market Street, 6th Floor
Philadelphia, PA 19104
USA

© 2026 Society for Industrial and Applied Mathematics

PDF

Help